

Mon Apr 7 07:36:48 2003

us-10-057-951-2.rpr

age 1

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 06:47:33 ; Search time 17.6747 Seconds
(without alignments)
1430.480 Million cell updates/sec

Title: US-10-057-951-2

Sequence: 1 MLAWQAFPLVSNMLAEAV.....PVDQSGSTPLMGQACTGCA 263

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: PIR.73.*

1: pirl.*

2: pirl.*

3: pirl.*

4: pirl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	134.5	11.1	562	1 UKRUT	t-plasminogen acti
2	134.5	11.1	359	1 A35029	t-plasminogen acti
3	134.5	11.1	359	1 A35029	t-plasminogen acti
4	146.2	10.6	255	1 A35029	hepatocyte growth
5	145.2	10.2	231	1 B3098	plasminogen acti
6	142	10.2	810	1 PIRU	plasminogen acti
7	138	9.9	169	2 A40522	plasminogen acti
8	138	9.9	433	1 UKRUT	macrophage-stimula
9	137	9.8	716	1 JC5061	u-plasminogen acti
10	135.5	9.7	442	1 UKRUT	u-plasminogen acti
11	134	9.6	431	2 JS0599	t-plasminogen acti
12	134	9.6	477	1 A34369	t-plasminogen acti
13	134	9.6	477	2 JS0598	macrophage-stimula
14	134	9.6	716	1 A40332	u-plasminogen acti
15	132	9.5	394	2 JS0600	u-plasminogen acti
16	132	9.5	431	1 UKRUT	plasminogen acti
17	130	9.3	810	2 A35029	plasminogen acti
18	128.5	9.2	433	1 A35005	u-plasminogen acti
19	128.5	9.2	433	1 A35005	u-plasminogen acti
20	126.2	9.1	433	1 A35005	u-plasminogen acti
21	125.2	9.1	433	1 A35005	u-plasminogen acti
22	124.5	8.9	810	2 B3098	hepatocyte growth
23	124	8.9	728	1 UKRUT	hepatocyte growth
24	124	8.9	728	1 UKRUT	hepatocyte growth
25	124	8.9	812	1 PIRU	plasminogen acti
26	123	8.8	593	2 A45281	coagulation factor
27	122.5	8.8	4548	1 S00657	apoptosis factor
28	122	8.8	603	2 S28941	coagulation factor
29	121.5	8.7	728	1 A35644	hepatocyte growth

30	120	8.6	433	1 UKRUT	u-plasminogen acti
31	120	8.6	790	1 PIRU	plasminogen acti
32	119.5	8.6	477	2 JS0597	hepatocyte growth
33	119	8.5	710	1 JS1283	hepatocyte growth
34	118.5	8.5	615	1 KFRH12	coagulation factor
35	118	8.5	432	1 S18932	u-plasminogen acti
36	117	8.4	625	1 TBBO	chromin (EC 3.4.2
37	116	8.3	1420	2 A32869	apoptosis factor
38	115.5	8.2	711	1 A47136	macrophage-stimula
39	114.5	8.2	812	1 PIRU	plasminogen acti
40	114	8.2	618	2 A35827	plasminogen acti
41	113	8.1	528	2 A35827	plasminogen acti
42	112	8.1	433	1 A35029	plasminogen acti
43	110.5	7.9	685	1 A48285	plasminogen acti
44	108.5	7.8	123	2 C61545	plasminogen acti
45	107	7.7	617	2 S10511	thrombin (EC 3.4.2

ALIGNMENTS

RESULT 1
UKRUT
t-plasminogen activator (EC 3.4.21.66) precursor [validated] - human
N/Alternate names: t-PA; tissue plasminogen activator
C/Species: Homo sapiens (man)
C/Date: 14-Nov-1983 #sequence revision 14-Nov-1983 #text change 08-Dec-2000
C/Accession: A94004; A23529; U70562; A93293; S02125; A91343; A93951; A91322; A54645; 160
R/Ref: Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984
A/Title: The structure of the human tissue-type plasminogen activator gene: correlation
A/Accession: A94004; A94004; M01D:6428137; F01D:6089198
A/Molecule type: DNA
A/Residues: 1-562 <NT>
A/Cross-references: GB:100141
A/Note: The codon given for residue 93 (ACC) is inconsistent with the authors' translation
R/Pfizer Degen, S.J.; Rajput, B.; Reich, E.
J. Biol. Chem. 261, 6972-6985, 1986
A/Title: The human tissue plasminogen activator gene.
A/Reference number: A23529; M01D:616143; F01D:3009482
A/Accession: A23529
A/Molecule type: DNA
A/Residues: 1-562 <DB>
A/Cross-references: GB:R03021; NID:G339817; PIR:AA98809.1; PID:G339818
R/Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuoka, S.; Higashio, K.
Agric. Biol. Chem. 55, 1225-1232, 1991
A/Title: Purification and characterization of tissue plasminogen activator secreted by
A/Reference number: U70562; M01D:31291340; F01D:1368681
A/Molecule type: mRNA
A/Residues: 31-562 <TR>
A/Cross-references: DDBJ:D01096; NID:G220128; PIR:BA00881.1; PID:G441174
A/Experimental source: embryonic lung fibroblast IMR-90 cells
A/Note: Part of this sequence, including the amino end of the mature protein, was conf.
R/Pennick, D.; Holmes, M.R.; Kohr, W.J.; Harkins, R.N.; Vohar, G.A.; Ward, C.A.; Bennett
Nature 301, 214-221, 1993
A/Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Escl
A/Reference number: A93293; M01D:8315262; F01D:6337343
A/Accession: A93293
A/Molecule type: mRNA
A/Residues: 1-562 <PEN>
A/Cross-references: GB:U00141
A/Experimental source: melanoma cells
R/Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M.
Nucleic Acids Res. 16, 5659, 1988
A/Title: The human tissue-type plasminogen activator cDNA from human
A/Reference number: S02125; M01D:86262579; F01D:3133640
A/Accession: S02125
A/Status: translation not shown
A/Molecule type: mRNA
A/Residues: 1-562 <SAS>
A/Cross-references: EMBL:X07933; NID:G37243; PIR:CAA30302.1; PID:G37244

A:Experimental source: fetal lung cells
A:Kagihara, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Okada, K.; Matsuda, K.
FEBS Lett. 189, 145-149, 1985
A:Title: Expression in *Escherichia coli* of finger-domain lacking tissue-type plasminogen activator cDNA
A:Reference number: A91343; MUID:85285620; PMID:3896853
A:Molecule type: mRNA
A:Accession: A91343A
A:Residues: 1-139 'G', 86-413 'E', 435-562 <ARG>
A:Experimental source: Detroit 562 cells; ATCC 138
A:Edlund, T.; NY, T.; Ranby, M.; Heden, L.O.; Palm, G.; Holmgren, E.; Josephson, S.
Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983
A:Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen activator
A:Reference number: A93951; MUID:83169656; PMID:6572897
A:Accession: A93951A
A:Molecule type: mRNA
A:Residues: 251-358 <EDL>
A:Experimental source: melanoma cells
A:Polh, G.; Kallstrom, M.; Bergsdorf, N.; Wallen, P.; Jorntvall, H.
Biochemistry 23, 3701-3707, 1984
A:Title: Isolation and characterization of a complementary DNA sequence for human tissue-type plasminogen activator
A:Reference number: A90489; MUID:85000468; PMID:6433976
A:Contents: annotation; melanoma cells; partial sequence of residues 36-562, active and inactive
A:Polh, G.; Kaplan, L.; Elmarsson, M.; Wallen, P.; Jorntvall, H.
FEBS Lett. 168, 29-32, 1984
A:Title: Differences between urokinase and melanoma forms of tissue plasminogen activator
A:Reference number: A91322; MUID:84189564; PMID:6538514
A:Accession: A91322A
A:Molecule type: protein
A:Residues: 33-45/311-320 <POH>
A:Experimental source: uterus
A:Note: in the uterus, cleavage of the activation peptide may also occur after 38-Gln
A:Van Dongenfeld, A.J.; Westma, H.; Bakkema, H.
Biochemistry 23, 3686-3689, 1984
A:Reference number: A37567; MUID:89033611; PMID:3021732
A:Contents: annotation; fibrin binding site
A:Verheijen, J.H.; Camper, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Bouwels, P.H.; Engerink, J.
EMBO J. 5, 3525-3530, 1986
A:Title: Involvement of finger domain and kringle 2 domain of tissue-type plasminogen activator in plasminogen activation
A:Reference number: A37568; MUID:8715171; PMID:3030730
A:Contents: annotation; fibrin binding site
A:Dodd, I.; Nunn, B.; Robinson, J.H.
Thromb. Haemost. 59, 523-528, 1986
A:Title: Isolation, identification and pharmacokinetic properties of human tissue-type plasminogen activator cDNA
A:Reference number: A60902; MUID:89044681; PMID:3142086
A:Contents: annotation; novel forms of expressed recombinant t-PA
A:Kawabe, T.; MacRae, R.; Patisson, J.; Bataillon, F.A.O.; Little, S.; Emtage, J.S.; Opdenacker, G.; Mutsaers, J.H.G.M.
Mol. Cell. Biochem. 125, 199-202, 1988
A:Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its expression in *Escherichia coli*
A:Reference number: A54645; MUID:86284200; PMID:3050401
A:Accession: A54645A
A:Molecule type: mRNA
A:Residues: 1-562 <HAR>
A:Cross-references: GB:M15518; NID:g190031; PIDN:AAA60111.1; PID:g190032
A:Note: parts of this sequence were confirmed by peptide sequencing
A:Reddy, V.B.; Garramone, A.J.; Sasek, H.; Wei, C.
DNA 6, 461-472, 1987
A:Title: Expression of human uterine tissue-type plasminogen activator in mouse cells using a plasmid vector
A:Reference number: I60110; MUID:88054470; PMID:2824147
A:Accession: I60110
A:Status: translated from GB/EMBL/DBD
A:Experimental source: melanoma cells
A:Residues: 1-562 <HAR>
A:Cross-references: GB:M18182; NID:g340176; PIDN:AAA36800.1; PID:g340177
A:Fishler, R.; Waller, E.K.; Grobel, G.; Thompson, D.; Tizabi, R.; Schleuning, W.D.
J. Biol. Chem. 260, 11233-11230, 1985
A:Title: Isolation and characterization of the human tissue-type plasminogen activator cDNA
A:Reference number: I55232; MUID:85289338; PMID:3161893
A:Accession: I55232A
A:Molecule type: DNA
A:Status: preliminary; translated from GB/EMBL/DBD
A:Residues: 1-36 <REZ>
A:Cross-references: GB:M11890; NID:g339837; PIDN:AAA61213.1; PID:g339839

[illegible]

F_309-559/Product: t-Plasminogen activator chain B #status predicted <BCH>
F_309-553/Domains: trypsin homology <TRY>
E_38-68,-66,-75,-83,-94,-88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297-344
E_149,-481/Binding site: carboxylate (Asn) (covalent) #status predicted
F_308-309/Cleavage site: Arg-ile (plasmin, trypsin) #status predicted
F_355,-404,510/Active site: His, Asp, Ser #status predicted

Query Match 11.1% Score 154; DB 1; Length 559;
Best Local Similarity 32.4%; Pred. No. 3,7e-05;
Matches 45; Conservative 14; Mismatches 54; Indels 26; Gaps 6;

Dc 25 CFWMGNHLYREDQTSPPAGRLCLMWLDASGLASPVS-----GAGNSHYCRNDE 75
Db 124 CFEQGQITGRGWSTAEINAGECIM--NSSALSQRKPYSAARRPAIKGLDNHMYCRNDR 181
Oy 76 DPRGPWCYSGEAGVEPKRPEDLRCE-----TTSQALPAPFTTEIGESSPG 124
Dc 182 DKR-FWCTVF-DAGRTTFETCPACRGPTEDCYVGSVTIRGHSHFI-SKASCLPW 257
Oy 125 ADPOGVAPANNAIPARSEA 143
Db 238 NSMILTKTYTMAMRANSQA 256

RESULT 3
A29941
C-plasminogen activator (EC 3.4.21.68) precursor - mouse
C/Species: Mus musculus (house mouse)
C/Date: 10-Sep-1999 #sequence-revision 10-Sep-1999 #text-1999
C/Accession: A29941; S48205; S48207; S48206
R/Rickles, R.J.; Darlow, A.L.; Strickland, S.
J. Biol. Chem. 263, 1563-1569, 1988
A/BioL: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mRNA
A/Reference number: A29941; PMID:88087303; PMID:2826484
A/Acession: A29941
A/Molecule type: cDNA
A/Residues: 1-559 <PIC>
R/Cross-references: GS:t03520, MID:g202109, PIRN:AAA0470.1, PID:g202110
R/LIJden, H.R., Vlasov, B.G., Beutin, V., Colten, D.
Biochim. Biophys. Acta 661-871, 1989
A/BioL: Characterization of the murine plasma fibrinolytic system.
A/Reference number: S48202; PMID:95010076; PMID:7523120
A/Acession: S48205
A/Molecule type: protein
A/Residues: 33-37, 'X', 39-40 <LIU>
A/Acession: S48207
A/Molecule type: protein
A/Residues: 309-16 <RI2>
A/Acession: S48206
A/Molecule type: protein
A/Residues: 33-37, 'X', 39-40 <LIW>
A/Superfamily: trypsin plasminogen activator; EGF homology; fibronectin type I repeat homologous; fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
E/F_1-17/Domains: signal sequence #status predicted <SIG>
F_18-28/Domains: propeptide #status predicted <PRO>
F_30-559/Product: t-Plasminogen activator #status predicted <MAT>
F_308-309/Product: t-Plasminogen activator chain A #status predicted <ACH>
F_38-75/Domains: fibronectin type I repeat homology <IFI>
F_34-166/Domains: EGF homology <EGF>
F_213-294/Domains: Kringle homology <KR1>
F_213-294/Domains: Kringle homology <KR2>
F_309-553/Domains: t-Plasminogen activator chain B #status predicted <BCH>
F_309-553/Product: t-Plasminogen activator chain B #status predicted <BCH>
F_38-68,-66,-75,-83,-94,-88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297-344
E_149,-481/Binding site: carboxylate (Asn) (covalent) #status predicted
F_308-309/Cleavage site: Arg-ile (plasmin, trypsin) #status predicted
F_355,-404,510/Active site: His, Asp, Ser #status predicted

Query Match 10.6% Score 147.5; DB 1; Length 559;
Best Local Similarity 37.0%; Pred. No. 0.00033;
Matches 34; Conservative 11; Mismatches 34; Indels 13; Gaps 4;

25 CFWDGNHLYREDQTSPPAGRLCLMWLDASGLASPVS-----GAGNSHYCRNDE 75

124 CPEBQ1TYRQETMTSGAGCINM--NSVSLVSLKPNARRRNMAIKIGLGNHNYCRBDD 181
125
126 DPRGPCVYSGAGVEKRPCEPEDIKCPETQ 107
127
128 DLK-PMCVYF-KAGKYTFECPSTPACPKGSE 211
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
9

A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-291 <SIE>
A:Cross-references: EMBL:X13097; NID:G35282; PIDD:CAA31489.1; PID:G35283
C:Comment: For the main splice form, see FIR:UKHUT. This form probably does not have pro
C:Genetics:
A:Gene: GDB:PLAT
A:Cross-references: GDB:119496; OMIM:173370
A:Map position: 8p12-8p12
A:Introns: 24/3, 39/1, 85/1, 122/1, 180/2, 211/1, 268/2
C:Superfamily: class plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: alternative splicing; fibrinolysis; glycoprotein; kringle
F:1-23/Domain: signal sequence #status predicted <PRO>
F:23-26/Domain: propeptide #status predicted <PRO>
F:26-36/Domain: plasminogen activator; repeat endothelial splice form #status pred
F:41-78/Domain: fibronectin type-III EGF
F:86-119/Domain: EGF homology <EGF>
F:127-208/Domain: kringle homology #status atypical <KR1>
F:215-291/Domain: kringle homology #status atypical <KR2>
F:41-71,69-78,86-97,91-108,110-119,127-208,148-190,179-203/Disulfide bonds: #status pred
Query Match 10.4%; Score 145.5; DB 2; Length 291;
Best Local Similarity 38.5%; Pred. No. 9.1e-05;
Matches 35; Conservative 6; Mismatches 37; Indels 13; Gaps 4;
QY 25 CFWDNQHLYREDQSPAPGLRLQMLDAQSGIASAPV-----GAGNHSYGRNDE 75
DB 127 CYDDQISYRGTSWSTAESGAECTNW--NSSALMAQMAVSGRPDAIRGLGNHYCRNPD 184
QY 76 DPRGWCYSGEAGVEKRPCEDLRCPETTS 106
DB 185 DSK-PWCTVP-KAGKTSSEFCSTPACSEGNS 213
RESULT 6
pHUT
plasmin (EC 3.4.21.7) precursor [validated] - human
N:Alternate names: plasminogen precursor [unsmet]
N:Contents: angiotensin; microplasma; plasminogen
C:Species: Homo sapiens (man)
C:Date: 24-Apr-1994 #sequence revision 02-Dec-1994 #text change 15-Sep-2000
C:Accession: A35229; 152242; A26646; 162738; 184609; S03735; A00929; A04627; A04625; A04
R:Peterson, T.E.; Martzen, M.R.; Ichinose, A.; Davie, E.W.
J. Biol. Chem. 265, 6104-6111, 1990
A:Title: Characterization of the gene for human plasminogen, a key proenzyme in the fibr
A:Reference number: A35229; MUID:90202879; PMID:2318848
A:Accession: A35229
A:Molecule type: DNA
A:Residues: 1-810 <PRT>
A:Cross-references: GB:05286; GB:M34276; NID:G19064; PIDD:AAA60113.1; PID:G387026
A:Experimental source: Leukoocyte; Lung fibroblast
A:Map position: 11p11-11p13
A:Keywords: Angiogenesis; Blood coagulation; Fibrinolysis; Hemostasis; Thrombolysis
Biochem. Biophys. Res Commun. 113, 118-120, 1980
A:Title: Definition of the transcriptional initiation site of human plasminogen gene in 11
A:Reference number: 152242; MUID:91097523; PMID:2266308
A:Accession: 152242
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-16 <MAL1>
A:Cross-references: GB:M6289; NID:G19092; PIDD:AAA36454.1; PID:G35363
R:Forsgren, M.; Raden, B.; Israelsson, M.; Larsson, K.; Heden, L.O.
FEBS Lett. 213, 254-260, 1987
A:Title: Molecular cloning and characterization of a full-length cDNA clone for human pl
A:Reference number: A26646; MUID:87162490; PMID:3030813
A:Accession: A26646
A:Molecule type: mRNA
A:Residues: 1-471, 'D', 473-810 <FOR>
A:Cross-references: GB:X05199; NID:G35530; PIDD:CAA28831.1; PID:G35531
A:Experimental source: Liver
A:Map position: 12p11-12p13
A:Keywords: Angiogenesis; Blood coagulation; Fibrinolysis; Hemostasis; Thrombolysis
Biochem. Biophys. Res Commun. 121, 1244-1250, 1984
A:Title: Characterization of a complementary deoxyribonucleic acid coding for human and
A:Reference number: 145961; MUID:05023311; PMID:6149561

A:Accession: 162738
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 292-471, 'D', 473-810 <MAL2>
A:Cross-references: GB:X02922; NID:G190112; PIDD:AAA60124.1; PID:G387031
A:Accession: 184609
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 367-419 <MAL3>
A:Cross-references: GB:X02921; NID:G190110; PIDD:AAA60123.1; PID:G190111
R:Brinholts, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Leigler, W.; Manneberg, M
Eur. J. Biochem. 114, 465-470, 1981
A:Title: Comparison of the primary structure of the N-terminal CNBR fragments of human,
A:Reference number: S03735; MUID:81212097; PMID:7238497
A:Accession: S03735
A:Molecule type: protein
A:Residues: 20-71, 'E', 73-76 <BRU>
R:Brinholts, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Leigler, W.; Manneberg, M
Submitted to the EMBO J. 1981
A:Reference number: A00929
A:Accession: A00929
A:Molecule type: protein
A:Residues: 20-71, 'E', 73-85, 87-106, 'D', 108-360, 'E', 362-810 <SOT>
R:Wiman, B.
Eur. J. Biochem. 76, 129-137, 1977
A:Title: Primary structure of the B-chain of human plasmin.
A:Reference number: A04627; MUID:77252545; PMID:142009
A:Accession: A04627
A:Molecule type: protein
A:Residues: 581-810 <WII>
R:Wiman, B.; Wallen, P.
Eur. J. Biochem. 50, 489-494, 1975
A:Title: Structural relationship between "glutamic acid" and "lysine" forms of human p
A:Reference number: A04625; MUID:75093329; PMID:122932
A:Accession: A04625
A:Molecule type: protein
A:Residues: 20-50, 'Q', 51-71, 'E', 73-85, 87-100 <W12>
R:Wiman, B.; Wallen, P.
Eur. J. Biochem. 58, 539-547, 1975
A:Title: Amino-acid sequence of the cyanogen-bromide fragment from human plasminogen ch
A:Reference number: A04626; MUID:76043692; PMID:126863
A:Accession: A04626
A:Molecule type: protein
A:Residues: 483-507, 'E', 509-604 <W13>
R:Robbins, K.C.; Bernabe, P.; Arzodon, L.; Summaria, L.
J. Biol. Chem. 248, 1631-1633, 1973
A:Title: The primary structure of human plasminogen. II. The histidine loop of human p
A:Reference number: A92125; MUID:73149248; PMID:4694729
A:Contents: annotation; active site
R:Roskopf, W.R.; Summaria, L.; Robbins, K.C.
J. Biol. Chem. 244, 3550-3557, 1969
A:Title: Studies on the active center of human plasmin. Partial amino acid sequence of
A:Reference number: A92046; MUID:69234739; PMID:4240117
A:Contents: annotation; active site
R:Roskopf, W.R.; Summaria, L.; Robbins, K.C.
J. Biol. Chem. 244, 3550-3557, 1969
A:Title: Structure of the omega-aminocarboxylic acid-binding sites of human plasminogen
A:Reference number: A92382; MUID:82213905; PMID:6919539
A:Contents: annotation; omega-aminocarboxylic acid binding sites
R:Wall, Z.; Patchy, L.
J. Biol. Chem. 259, 13690-13694, 1984
A:Title: The fibrin-binding site of human plasminogen. Arginines 32 and 34 are essential
A:Reference number: A92458; MUID:85054794; PMID:6094526
A:Contents: annotation; fibrin binding site; omega-aminocarboxylic acid binding site
R:Caio, Y.; Yi, R.W.; Davidson, D.; Schaller, J.; Marti, D.; Soehndel, S.; McCance, S.G.
J. Biol. Chem. 271, 29461-29467, 1996
A:Title: Kringle domains of human angiotensin. Characterization of the anti-proliferative
A:Reference number: A58811; MUID:97067211; PMID:8910613
A:Contents: annotation
R:Liijnen, H.R.; Ugwu, F.; Bani, A.; Collen, D.
Biochemistry 37, 4699-4702, 1998
A:Title: Generation of an angiotensin-like fragment from plasminogen by stromelysin-1 (

A:Reference number: A58812; PMID:9548733; PMID:9548733
A:Contents: annotation
R:Tullinsky, A.; Mulichak, A.M.
submitted to the Brookhaven Protein Data Bank, July 1991
A:Reference number: A51341; PDB:1PK4
A:Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 376-454
R:Tullinsky, A.; Wu, T.P.
submitted to the Brookhaven Protein Data Bank, July 1991
A:Reference number: A51498; PDB:2PK4
A:Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 377-454
R:Wu, T.P.; Tullinsky, A.
submitted to the Brookhaven Protein Data Bank, August 1993
A:Reference number: A51911; PDB:1PKR
A:Contents: annotation; X-ray crystallography, 2.48 angstroms, residues 102-181
R:Padmanabhan, K.; Tullinsky, A.
submitted to the Brookhaven Protein Data Bank, April 1994
A:Reference number: A52408; PDB:1PMK
A:Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 377-454
R:Tullinsky, A.; Mathews, I.I.
submitted to the Brookhaven Protein Data Bank, December 1995
A:Reference number: A65324; PDB:1CEA
A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
R:Tullinsky, A.; Mathews, I.I.
submitted to the Brookhaven Protein Data Bank, December 1995
A:Reference number: A65325; PDB:1CEB
A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
R:Mulichak, A.M.; Tullinsky, A.; Ravichandran, K.G.
Biochemistry 30, 10576-10588, 1991
A:Title: Crystal and molecular structure of human plasminogen kringle 4 refined at 1.9 Å
A:Reference number: A58819; PMID:92031502; PMID:1657148
A:Contents: annotation
R:Wu, T.P.; Padmanabhan, K.; Tullinsky, A.; Mulichak, A.M.
Biochemistry 30, 10589-10594, 1991
A:Title: The refined structure of the epsilon-aminocaproic acid complex of human Plasmin
A:Reference number: A58818; PMID:92031503; PMID:1657149
A:Contents: annotation
R:de Vos, A.M.; Olcese, M.H.; Kelley, R.F.; Padmanabhan, K.; Tullinsky, A.; Westbrook, M.
Biochemistry 31, 270-279, 1992
A:Title: Crystal structure of the kringle 2 domain of tissue plasminogen activator at 2 Å
A:Reference number: A53485; PMID:9211803; PMID:1110013
A:Contents: annotation; X-ray crystallography, 2.4 angstroms
R:Sec, A.; Tesch, M.; Wüthrich, K.; Yanagisawa, M.; Yanagisawa, A.
Reference number: A65590; PDB:1KRN
submitted to the Brookhaven Protein Data Bank, June 1995
R:Reichman, M.; Llinas, M.
submitted to the Brookhaven Protein Data Bank, August 1996
A:Reference number: A65803; PDB:1HPJ
A:Contents: annotation; conformation by (1)H-NMR, residues 103-181
R:Rejzender, M.; Llinas, M.
submitted to the Brookhaven Protein Data Bank, August 1996
A:Reference number: A65804; PDB:1HPK
A:Contents: annotation; conformation by (1)H-NMR, residues 103-181
R:Rejzender, M.R.; Llinas, M.
Biochemistry 221, 927-937, 1994
A:Title: (1)H-NMR assignments and secondary structure of human plasminogen kringle 1.
A:Reference number: S43645; PMID:9427157; PMID:8181475
A:Contents: annotation; conformation by (1)H-NMR, residues 96-184
R:Rejzender, M.R.; Llinas, M.
Eur. J. Biochem. 221, 939-949, 1994
A:Title: Solution structure of the epsilon-aminohexanoic acid complex of human plasminogen
A:Reference number: A58817; PMID:9427158; PMID:8181476
A:Contents: annotation; conformation by (1)H-NMR
C:Comment: plasminogen is synthesized by the kidney and is present in plasma and many of
d PIR:1YHGBA).
C:Comment: Plasmin is inactivated by alpha-2-antiplasmin (see PIR:1YHUA2) immediately after
rg-880, resulting in two chains connected by two disulfide bonds. Without the inhibitor
C:Comment: Microplasmin is formed by autocatalytic cleavage of plasmin under artificial con
C:Comment: Streptelysin I (see PIR:1KHUS1) acts on plasminogen to produce angiotensin. T
ing solid tumors.
A:Gene: GDB:PLG
A:Genetics:

```

A:Cross-references: GDB:119498; OMIM:113350
A:Map position: 6q26-qc27
A:Introns: 17/1; 62/2; 98/1; 136/2; 183/1; 223/2; 263/1; 317/2; 366/1; 419/2; 480/1; 523/2
C:Function:
A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a variety of tissues; cleaves the walls of the granatic follicle; also activates the urokinase-type plasminogen activator
A:Pathway: fibrinolysis
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: angiogenesis inhibitor; blood; duplication; fibrinolysis; glycoprotein; hydrolase
F:1-96/Domain: plasminogen-related protein precursor homology <PLH>
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-910/Product: plasminogen #status experimental <PRO>
F:20-96/Domain: activation peptide #status experimental <APT>
F:79-466/Product: angiotensin peptide #status experimental <AST>
F:97-580, 581-810/Product: plasmin #status experimental <MA>
F:97-580/Domain: plasmin chain A #status experimental <CHA>
F:103-181/Domain: kringle homology <KR1>
F:185-362/Domain: kringle homology <KR2>
F:217-352/Domain: kringle homology <KR3>
F:317-454/Domain: kringle homology <KR4>
F:481-580/Domain: kringle homology <KR5>
F:550-580, 581-810/Product: microplasmin #status experimental <MMT>

Query Match          10.2%; Score 142; DB 1; Length 810;
Best Local Similarity 37.2%; Pred. No. 0.00058;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

Ox 25 CFWNDGHLVREDTSPADGLRCLNWLDA-----QGSLASAPVSGAGNHSYCRNDDPPG 79
Db 103 CXTGNGKRYKMTGSKTKNGITCKQKMSSTPRHPRSPRHPSGL-ENNYCRNDDNDPPG 161
Ox 80 PWCYVSGEAGVPEKR--PCEDIRCPCE
Db 162 PWCITTD----PKRIVYCDLICEE 183

RESULT 7
A:00522
A:Plasmin (EC 3.4.21.7) precursor - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 28-Feb-1992 #sequence_revision 17-Apr-1993 #ext_change 16-Jul-1999
C:Accession: J.J. Makker, S.P. R.Kanaiya, J.J.
A:1801 Chem. 266, 10825-10829, 1991
A:Title: Identification of the rat Heymann nephritis autoantigen (GP330) as a precursor
A:Reference number: A40522; MUID:91250378; PMID:1645711
A:Accession: A40522
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-169 <XN>
A:Cross-references: GB:M6283; NID:926215; PID:AAA1884.1; PID:G554488
A>Note: the authors translated the codon TCT for residue 76 as Ala
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:34-112/Domain: kringle homology <KR>
F:34-112,55-95,83-107/Disulfide bonds: #status predicted

Query Match          9.9%; Score 138; DB 2; Length 169;
Best Local Similarity 31.8%; Pred. No. 0.00021;
Matches 35; Conservative 13; Mismatches 48; Indels 14; Gaps 5;

Ox 25 CFWNDGHLVREDTSPADGLRCLNWLDA-----QGSLASAPVSGAGNHSYCRNDDPPG 79
Db 34 CYQNGKSKYRSTSTNTTKKQSGWVSMTPHSHSKTPANFPDGL-ENNYCRNDDNDG 92
Ox 80 PWCYVSGEAGVPEKR--PCEDIRCPCEFTTSQALPAFTTICDASBEGADE 127
Db 93 PWCFTTD---PSVFWCYCNLKRCSFTGGV--ASDAIVDPVPAEPGIS 136

RESULT 8
A:plasminogen activator (EC 3.4.21.73) precursor - yellow baboon

```

C.Species: Papio cynocephalus, Papio hamadryas cynocephalus (Yellow baboon)
C.Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Jun-1999
C.Accession: S14687, S08651
R.Au: Y.P.T.; Wang, T.W.; Clower, A.W.
Nucleic Acids Res 18, 3411, 1990
A.Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminogen
A.Reference number: S14687, MUID:90287234; PMID:2113276
A.Accession: S14687
A.Molecule type: mRNA
A.Residues: 1-433 <AB>
A.Cross-references: EMBL:X51935; NID:G38130; PDB:CAA6200.1; PID:G38131
C.Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C.Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F.1-20/Domain: signal sequence #status predicted <SIG>
F.21-176/Product: plasminogen activator chain A #status predicted <ACH>
F.30-61/Domain: EGF homology <EGF>
F.69-150/Domain: kringle homology <KRG>
F.178-433/Product: plasminogen activator chain B #status predicted <BCH>
F.178-431/Domain: trypsin homology <TRY>
F.167-298,208-224,216-287,315-384,347-363,374-402/Disulfide bonds: #status predicted
F.223,274,378/Active site: His, Asp, Ser #status predicted
F.324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.9%; Score 138; DB 1; Length 433;
Best Local Similarity 32.7%; Pred. No. 0.00061;
Matches 32; Conservative 14; Mismatches 36; Indels 16; Gaps 4;

Qy 25 CFMDNGHLYREDQTSPPAGRLCLNWLDA-----QSGIAPVSGAGNHSYCRNPEDP 77
Db 69 CYGNGHFRGKASTDTMGSCIAVNSATVLTQYHARSDALDGLGKHNCRNP-NR 127

Qy 78 RGPWCYVSGAGVPEK-----RPECEDLRCPETTSQAL 109
Db 128 RRPWCYV--QVGLKRVQECVWVNCADKSSPPEEL 163

RESULT 9
JC5061
C.Species: Rattus norvegicus (Norway rat)
C.Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 16-Jun-2000
C.Accession: JC5061
R.Ohshiro, K.; Iwama, A.; Matsuno, K.; Esaki, T.; Sakamoto, O.; Hamaguchi, I.; Takasu, N
Biochem. Biophys. Res. Commun. 227, 273-280, 1996
A.Title: Molecular cloning of rat urokinase-stimulating protein and its involvement in
A.Reference number: JC5061; MUID:97011126; PMID:8858136
A.Molecule type: mRNA
A.Residues: 1-716 <CH>
A.Cross-references: EMBL:X5096; NID:G1669718; PDB:CAA64473.1; PID:G1669719
C.Superfamily: disulfide-bonded heterodimer of chains derived from the same precursor
C.Keywords: duplication; glycoprotein; growth factor; kringle
F.1-31/Domain: signal sequence #status predicted <SIG>
F.32-488,489-716/Product: macrophage-stimulating protein 1 #status predicted <MAT>
F.32-488/Domain: macrophage-stimulating protein 1 alpha chain #status predicted <ACH>
F.310-186/Domain: kringle homology <KR12>
F.191-268/Domain: kringle homology <KR13>
F.292-370/Domain: kringle homology <KR14>
F.379-457/Domain: kringle homology <KR15>
F.489-716/Domain: macrophage-stimulating protein 1 beta chain #status predicted <BCH>
F.489-709/Domain: trypsin homology <TRY>
F.712,305,620/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.8%; Score 137; DB 1; Length 716;
Best Local Similarity 27.3%; Pred. No. 0.0013;
Matches 48; Conservative 12; Mismatches 56; Indels 60; Gaps 8;

Qy 2 LIAVQAFLVSNMLLAEEYSG-----GCFMDNGHLYREDQTSPPAGRLCLNWL 49
Db 80 LIPWTC-----HSLAACHHSHSLCDLFOKDYVTCIMNGASRYGVARTADGHPQAW 134

Qy 50 ---LDAOSGIAPVSGAGNHSYCRNPEDPDRGPWCYV----- 85

Db 135 SRAFPNDKHTPTKNGL--EENFCRNPDDPDRPWCYTTNNSVRFQSGIKSCREAVCVW 193

Qy 86 -----GAGVPEK-RPCE--DLRCPET-----TSQALPAFTTEIOESBGP 123
Db 194 CNGEDYRGSEVDTESGREGCQRMWDLQHPHSHPFHPEKFPDKALKNYCRNPASERP 249

RESULT 10
UKRG
u-plasminogen activator (EC 3.4.21.73) precursor - pig
N.Alternate names: uPA
C.Species: Sus scrofa domestica (domestic pig)
C.Date: 04-Dec-1984 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998
C.Accession: A00932
R.Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.
Nucleic Acids Res. 12, 9525-9541, 1984
A.Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.
A.Reference number: A00932; MUID:85087954; PMID:6096832
A.Accession: A00932
A.Molecule type: DNA
A.Residues: 1-240, 'H', 242-442 <NAG1>
A.Experimental source: kidney cell line LLC-PK1
R.Nagamine, Y.
Submitted to the Protein Sequence Database, December 1986
A.Reference number: A37566
A.Contents: annotation; correction to residue 241
C.Intentions: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 268/1; 335/1; 384/3
C.Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology.
C.Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F.1-20/Domain: signal sequence #status predicted <SIG>
F.21-186/Product: urokinase-type plasminogen activator chain A #status predicted <AC
F.23-64/Domain: EGF homology <EGF>
F.72-153/Domain: kringle homology <KRG>
F.130-442/Product: plasminogen activator chain B #status predicted <B
F.130-442/Domain: trypsin homology <TRY>
F.152/450/Domain: trypsin homology <TRY> (covalent) #status predicted
F.122,329,326,428,229,324,393,356-372,383-411/Disulfide bonds: #status predicted
F.433,286,387/Active site: His, Asp, Ser #status predicted

Query Match 9.7%; Score 135.5; DB 1; Length 442;
Best Local Similarity 36.9%; Pred. No. 0.001;
Matches 38; Conservative 7; Mismatches 37; Indels 21; Gaps 5;

Qy 25 CFMDNGHLYREDQTSPPAGRLCLNWLDAQSGT---ASAPVS---GAGNHSYCRNPEDP 77
Db 72 CFEGNGHFRGKASTDTMGSCIAVNSATVLTQYHARSDALDGLGKHNCRNP-NQ 130

Qy 78 RGPWCYVSGAGVPEK-----RPECEDLRCPETTSQ 107
Db 131 RRPWCYVQGLKRVQECVWVNCADKSSPPEEL 173

RESULT 11
JS0599
t-plasminogen activator (EC 3.4.21.68) beta precursor - common vampire bat
N.Alternate names: tissue plasminogen activator
C.Species: Desmodus rotundus (common vampire bat)
C.Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jun-1999
C.Accession: JS0599
R.Kratzschmar, J.; Heandler, B.; Langer, G.; Botdel, W.; Briggmann, P.; Alagon, A.; Do
Gene 105, 229-237, 1991
A.Title: The plasminogen activator family from the salivary gland of the vampire bat De
A.Reference number: JS0597; MUID:92039036; PMID:1937019
A.Accession: JS0599
A.Molecule type: mRNA
A.Residues: 1-431 <KRA>
A.Cross-references: GB:M63989; NID:G166076; PDB:AAA31594.1; PID:G166077
C.Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat ho
C.Keywords: fibrinolytic; glycoprotein; hydrolase; kringle; serine proteinase
F.1-21/Domain: signal sequence #status predicted <SIG>
F.22-36/Domain: propeptide #status predicted <PRO>

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 06:30:08 ; Search time 9.89785 Seconds
(without alignments)

1102.085 Million cell updates/sec

Title: US-10-057-951-2

Perfect score: 1393

Sequence: 1 MLAWQAFVSNMLAEAY.....PVDPGSGTPLMGQATPGGA 263

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_40*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	ID	Description
1	160.5	11.5	566	1 TPA_BOVIN	Q28198 bos taurus
2	154.5	11.1	562	1 TPA_HUMAN	P00750 homo sapien
3	154	11.1	559	1 TPA_RAT	P19637 rattus norv
4	150	10.8	653	1 HGF_MOUSE	Q9R098 mus musculu
5	147.5	10.6	559	1 TPA_MOUSE	P11214 mus musculu
6	146.5	10.5	655	1 HGF_HUMAN	Q04756 homo sapien
7	142	10.2	810	1 PLMN_HUMAN	P00747 homo sapien
8	138	9.9	169	1 PLMN_RAT	Q01177 rattus norv
9	138	9.9	433	1 UROK_PAPCY	P16227 papio cynoc
10	135.5	9.7	442	1 UROK_PIG	P04185 sus scrofa
11	134	9.6	431	1 UROK_DESRO	P98121 desmodus ro
12	134	9.6	477	1 UROK_DESRO	P15638 desmodus ro
13	134	9.6	716	1 HGF_MOUSE	P26928 mus musculu
14	132	9.5	394	1 UROK_DESRO	P49150 desmodus ro
15	132	9.5	431	1 UROK_HUMAN	P00740 homo sapien
16	130	9.3	810	1 PLMN_RAT	Q29485 erinaceus e
17	128.5	9.2	434	1 UROK_CHICK	P15120 gallus gall
18	127	9.1	728	1 UROK_BOVIN	Q05889 bos taurus
19	126.5	8.9	810	1 HGF_MOUSE	Q08048 mus musculu
20	124.5	8.9	728	1 PLMN_MACMU	P12445 macaca mula
21	124	8.9	622	1 THRB_HUMAN	P00734 homo sapien
22	124	8.9	728	1 HGF_HUMAN	P14210 homo sapien
23	123	8.8	812	1 PLMN_BOVIN	P06868 bos taurus
24	123	8.8	593	1 PAI2_BOVIN	P98140 bos taurus
25	122.5	8.8	4548	1 APOA_HUMAN	P08519 homo sapien
26	122	8.8	603	1 PAI2_CAVPO	Q04962 cavia porce
27	121.5	8.7	728	1 HGF_RAT	P17945 rattus norv
28	120.5	8.7	333	1 PLMN_CANFA	P80009 canis famli
29	120	8.6	433	1 UROK_MOUSE	P06867 sus scrofa
30	120	8.6	790	1 PLMN_PIG	P06867 sus scrofa
31	119.5	8.5	477	1 UROK_DESRO	P98148 desmodus ro
32	118.5	8.5	615	1 PAI2_HUMAN	P00748 homo sapien
33	118	8.5	432	1 UROK_RAT	P29598 rattus norv

34	117	8.4	625	1 THRB_BOVIN	P00735 bos taurus
35	116	8.3	1420	1 APOA_MACMU	P14417 macaca mula
36	115.5	8.3	711	1 HGF_HUMAN	P26927 homo sapien
37	114.5	8.2	812	1 PLMN_MOUSE	P09918 mus musculu
38	114	8.2	473	1 KREM_MOUSE	Q92943 mus musculu
39	114	8.2	473	1 KREM_RAT	Q92484 rattus norv
40	114	8.2	475	1 KREM_HUMAN	Q96m8 homo sapien
41	114	8.2	618	1 THRB_MOUSE	P19221 mus musculu
42	109.5	7.9	1709	1 SN_HUMAN	Q9bzz2 homo sapien
43	107	7.7	617	1 THRB_RAT	P18292 rattus norv
44	103.5	7.4	343	1 PLMN_SHEEP	P81286 ovis aries
45	101.5	7.3	325	1 PLMN_PETMA	P33574 petromyzon

ALIGNMENTS

```

RESULT 1
ID   TPA_BOVIN          STANDARD;          PRT;          566 AA.
AC   Q28198
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE   (t-PA) (t-plasminogen activator).
GN   Bos taurus (Bovine).
OS   Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OX   Bovidae; Bovinae; Bos.
RN   NCBI_TaxID=9913;
RP   [1]
RP   SEQUENCE FROM N.A.
RC   Tissue-Kidney;
RA   Rayn P. Berglund L., Petersen T.E.;
RT   "Cloning and characterization of the bovine plasminogen activators uPA
and tPA."
RL   Int. Dairy J. 5:605-617(1995).
CC   -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZIMMOEN PLASMINOGEN
CC   TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC   CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC   ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC   MANY OTHER PHYSIOPATHOLOGICAL EVENTS.
CC   -!- CATALYTIC ACTIVITY: Specific cleavage of Arg|-Val bond in
CC   plasminogen to form plasmin.
CC   -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC   BOND.
CC   -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC   -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC   PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC   ARG-314 CATALYZED BY PLASMIN. TISSUE KALLIKREIN OR FACTOR XA.
CC   -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC   CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC   -!- SIMILARITY: BELONGS TO TRYPSIN-LIKE FAMILY S1.
CC   -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE 1 DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC   -----
CC   This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; X65800; CAAS9735.1; -.
DR   HSSP; P00750; 1RTF.
DR   MEROPS; S01.232; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR000083; Fibrinctn1.

```


RA Schleuning W.-D.;
 RT "Isolation and characterization of the human tissue-type plasminogen
 RT activator structural gene including its 5' flanking region.";
 RL J. Biol. Chem. 260:11223-11230(1985).
 RN [9]
 RP SEQUENCE OF 31-562 FROM N.A.
 RX MEDLINE=91291340; PubMed=1366681;
 RA Itagaki Y., Yasuda H., Morinaga T., Mitsuda S., Higashio K.;
 RT "Purification and characterization of tissue plasminogen activator
 RT secreted by human embryonic lung diploid fibroblasts, IMR-90 cells".
 RL Agric. Biol. Chem. 55:1225-1232(1991).
 RN [10]
 RP SEQUENCE OF 36-562.
 RC TISSUE-Melanoma;
 RX MEDLINE=85000468; PubMed=6433976;
 RA Pohl G., Kaelin-Lang A., Bergsdorf N., Wallen P., Joernvall H.;
 RT "Tissue plasminogen activator: peptide analyses confirm an indirectly
 RT derived amino acid sequence, identify the active site serine residue,
 RT establish glycosylation sites, and localize variant differences.";
 RL Biochemistry 23:3701-3707(1984).
 RN [11]
 RP SEQUENCE OF 33-52 AND 311-330.
 RC TISSUE-Melanoma;
 RX MEDLINE=83209620; PubMed=6682760;
 RA Wallen P., Pohl G., Bergsdorf N., Raahy M., Ny T., Joernvall H.;
 RT "Purification and characterization of a melanoma cell plasminogen
 RT activator.";
 RL Eur. J. Biochem. 132:681-686(1983).
 RN [12]
 RP SEQUENCE FROM N.A. (SHORT ISOFORM).
 RC TISSUE-Umbilical vein;
 RX MEDLINE=90192129; PubMed=2107528;
 RA Siebert P.D., Fong K.;
 RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from
 RT human endothelial cells.";
 RL Nucleic Acids Res. 18:1086-1086(1990).
 RN [13]
 RP STRUCTURE OF CARBOHYDRATES.
 RX MEDLINE=90092112; PubMed=2513186;
 RA Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.;
 RT "Carbohydrate structure of recombinant human uterine tissue
 RT plasminogen activator expressed in mouse epithelial cells";
 RL Eur. J. Biochem. 186:273-286(1989).
 RN [14]
 RP CARBOHYDRATE-LINKAGE SITE THR-96.
 RX MEDLINE=91159408; PubMed=1900431;
 RA Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.;
 RT "Tissue plasminogen activator has an O-linked fucose attached to
 RT threonine-61 in the epidermal growth factor domain.";
 RL Biochemistry 30:2311-2314(1991).
 RN [15]
 RP DISULFIDE BONDS IN KRINGLE 2.
 RX MEDLINE=91244765; PubMed=1645336;
 RA Vlahos C.J., Wilhelm O.G., Hassell T., Jaskunas S.R., Bang N.U.;
 RT "Disulfide pairing of the recombinant kringle-2 domain of tissue
 RT plasminogen activator produced in *Scherichia coli*.";
 RL J. Biol. Chem. 266:10070-10072(1991).
 RN [16]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF CATALYTIC DOMAIN.
 RX MEDLINE=96200985; PubMed=8613982;
 RA Lamba D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,
 RT "The 2.3 A crystal structure of the catalytic domain of recombinant
 RT two-chain human tissue-type plasminogen activator.";
 RL J. Mol. Biol. 258:117-135(1996).
 RN [17]
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.
 RX MEDLINE=97449126; PubMed=9305622;
 RA Renatus M., Engh R.A., Stubbs M.T., Huber R., Fischer S., Kohnert U.,
 RT "Lysine 156 promotes the anomalous proenzyme activity of tPA: X-ray
 RT crystal structure of single-chain human tPA.";
 RL EMBO J. 16:4797-4805(1997).

RN [18]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF KRINGLE 2.
 RX MEDLINE=92118803; PubMed=1310033;
 RA De Vos A., Ultsch M.H., Kelley R.F., Padmanabhan K., Tulinsky A.,
 RA Westbrook M.L., Kosiakof A.A.;
 RT "Crystal structure of the kringle 2 domain of tissue plasminogen
 RT activator at 2.4-A resolution.";
 RL Biochemistry 31:270-279(1992).
 RN [19]
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=90122799; PubMed=2558718;
 RA Byeon I.-J., Kelley R.F., Llinas M.;
 RT "1H NMR structural characterization of a recombinant kringle 2 domain
 RT from human tissue-type plasminogen activator.";
 RL Biochemistry 28:9350-9360(1989).
 RN [20]
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=9120042; PubMed=1901789;
 RA Byeon I.-J., Kelley R.F., Llinas M.;
 RT "Kringle-2 domain of the tissue-type plasminogen activator. 1H-NMR
 RT assignments and secondary structure.";
 RL Eur. J. Biochem. 197:155-165(1991).
 RN [21]
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=92106329; PubMed=1762144;
 RA Byeon I.-J., Llinas M.;
 RT "Solution structure of the tissue-type plasminogen activator kringle
 RT 2 domain complexed to 6-aminohexanoic acid an antifibrinolytic
 RT drug.";
 RL J. Mol. Biol. 222:1035-1051(1991).
 RN [22]
 RP STRUCTURE BY NMR OF 38-85.
 RX MEDLINE=92292163; PubMed=1602484;
 RA Downing A.K., Driscoll P.C., Harvey T.S., Dudgeon T.J., Smith B.O.,
 RA Bacon M., Campbell I.D.;
 RT "The solution structure and backbone dynamics of the fibronectin type
 RT I and epidermal growth factor-like pair of modules of tissue-type
 RT plasminogen activator.";
 RL J. Mol. Biol. 225:821-833(1992).
 RN [23]
 RP STRUCTURE BY NMR OF 36-126.
 RX MEDLINE=9607104; PubMed=7582899;
 RA Smith B.O., Downing A.K., Driscoll P.C., Dudgeon T.J., Campbell I.D.;
 RT "The solution structure and backbone dynamics of the fibronectin type
 RT I and epidermal growth factor-like pair of modules of tissue-type
 RT plasminogen activator.";
 RL Structure 3:823-833(1995).
 CC -I- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN
 CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
 CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
 CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
 CC MANY OTHER PHYSIOLOGICAL EVENTS.
 CC -I- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -I- SUBUNIT: BINDS TO FIBRIN WITH HIGH AFFINITY. THIS INTERACTION
 CC

Query Match 11.1%; Score 154.5; DB 1; Length 562;
 Best Local Similarity 39.6%; Pred. No. 1.1e-05;
 Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 25 CFMNGHLYVEDDTSPAPGRCLNWLDAQSLASAPVS-----GAGHSYCRNPDE 75
 DB 127 CYDGGISYNGTSTVDSAGACTNW--NSSALAQKPYSGRRPDALRLGLGNHNYCRNPDR 184
 QY 76 DPRGPWCYVSGEAGVPEKRPCEDLRCPETTS 106
 DB 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 213

RESULT 3
 TPA_RAT
 ID_TPA_RAT STANDARD; PRT; 559 AA.
 AC P19637;

DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (rPA)
DE (c-pa) (t-plasminogen activator).
GN PLAT.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89170114; PubMed=3148445;
RA NY T., Leonardson G., Heuvel A.J.W.;
RT "Cloning and characterization of a cDNA for rat tissue-type
RT plasminogen activator";
RL DNA 7:671-677(1988).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=90130448; PubMed=2105315;
RA Feng P., Ohlsson M., Ny T.;
RT "The structure of the TATA-less rat tissue-type plasminogen activator
RT gene. Specific sequence divergences in the promoter predict
RT differences in regulation of gene expression.";
RL J. Biol. Chem. 265:2022-2027(1990).
CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOMEN PLASMINOGEN
CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC MANY OTHER PHYSIOLOGICAL EVENTS.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC BOND.
CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 BGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL, M23697; AAA41812.1; -
DR EMBL, M31197; AAA42261.1; -
DR EMBL, M31185; AAA42261.1; JOINED.
DR EMBL, M31186; AAA42261.1; JOINED.
DR EMBL, M31187; AAA42261.1; JOINED.
DR EMBL, M31188; AAA42261.1; JOINED.
DR EMBL, M31189; AAA42261.1; JOINED.
DR EMBL, M31190; AAA42261.1; JOINED.
DR EMBL, M31191; AAA42261.1; JOINED.
DR EMBL, M31192; AAA42261.1; JOINED.
DR EMBL, M31193; AAA42261.1; JOINED.
DR EMBL, M31194; AAA42261.1; JOINED.
DR EMBL, M31195; AAA42261.1; JOINED.
DR EMBL, M31196; AAA42261.1; JOINED.
DR EMBL, A19618; CAA01482.1; -
DR PIR, A31597; A31597.
DR HSSP, P00750; IRTF.
DR MEROPS, S01.232; -
DR InterPro, IPR001314; Chymotrypsin.

DR InterPro, IPR000561; EGF-like.
DR InterPro, IPR000083; Fibrinctn.
DR InterPro, IPR000001; Kringle.
DR InterPro, IPR01254; Ser_protease_Try.
DR Pfam, PF00008; EGF, 1.
DR Pfam, PF00039; fnl, 1.
DR Pfam, PF00051; kringle, 2.
DR Pfam, PF00089; trypsin, 1.
DR PRINTS, PR00722; CHYMOTRYPSIN.
DR PRINTS, PR00018; KRINGLE.
DR ProDom, PD00395; kringle, 2.
DR SMART, SM00181; EGF, 1.
DR SMART, SM00058; FN1, 1.
DR SMART, SM00130; KR, 2.
DR SMART, SM00020; Tryp_Spc, 1.
DR PROSITE, PS00022; EGF_1, 1.
DR PROSITE, PS01186; EGF_2, 1.
DR PROSITE, PS01253; FIBRONECTIN_1, 1.
DR PROSITE, PS00021; KRINGLE_1, 2.
DR PROSITE, PS00070; KRINGLE_2, 2.
DR PROSITE, PS00240; TRYPSIN_DOM, 1.
DR PROSITE, PS00134; TRYPSIN_HIS, 1.
DR PROSITE, PS00135; TRYPSIN_SER, 1.
KW plasminogen activator; Hydrolase; Serine protease; Glycoprotein;
KW plasma; kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 17
FT PROPEP 18 29
FT CHAIN 30 559
FT CHAIN 30 308
FT CHAIN 309 559
FT DOMAIN 36 78
FT DOMAIN 79 117
FT DOMAIN 124 205
FT DOMAIN 213 294
FT DOMAIN 309 559
FT ACT SITE 355 355
FT ACT SITE 404 404
FT ACT SITE 510 510
FT DISULFID 38 68
FT DISULFID 66 75
FT DISULFID 83 94
FT DISULFID 88 105
FT DISULFID 107 116
FT DISULFID 124 205
FT DISULFID 145 187
FT DISULFID 176 200
FT DISULFID 213 294
FT DISULFID 234 276
FT DISULFID 265 289
FT DISULFID 297 428
FT DISULFID 340 356
FT DISULFID 348 417
FT DISULFID 442 516
FT DISULFID 474 490
FT DISULFID 506 534
FT CARBOHYD 149 149
FT CARBOHYD 481 481
FT CONFLICT 380 380
SQ SEQUENCE 559 AA; 62903 MW; 7DBD1809C1DC21 CRC64;
Query Match 11.1%; Score 154; DB 1; Length 559;
Best Local Similarity 32.4%; Pred. No. 1, 2e-05;
Matches 45; Conservative 14; Mismatches 54; Indels 26; Gaps 6;
OY 25 CPMDNGHLVREDQTSAPRLCLNWLDAQSLGASAPV-----GAGNHSYCNPDE 75
DB 124 CFEQGQITTRGWSHTEAENGACINW--NSSLBSQKPSARSAPNAIKGLGNHNYCNPNR 181
OY 76 DPRGWCYVSGEAGVPEKPCEDLRCPE-----TTSQLPAFTTEIQEASEGFG 124
DB 182 DVK-PWCYVF-KAGKYTEFCSTPACPKGPTEDCVGKGVTRGTHFTT--SKASCLPW 237

QY 125 ADEVOVFAPAMALPARSEA 143
DB 238 NSMILIGKTYTAMRANSOA 256

RESULT 4
HGFA_MOUSE
ID HGFA_MOUSE STANDARD; PRT; 653 AA.
AC Q9R098; Q90KV4; 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (BC 3.4.21.-) (HGF activator) (HGFA).
GN HGFA_MOUSE
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/c;
RA Itoh H., Kataoka H., Koono H.;
RT "Mouse hepatocyte growth factor activator."
RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RA van Adelsberg J.S., Sehgal S., Kukes A., Brady C., Barasch J., Yang J., Huan Y.;
RT "Activation of HGF by endogenous HGF activator is required for metanephric kidney morphogenesis in vitro."
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY CONVERTING IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM (BY SIMILARITY).
CC -1- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE II DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by, and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF099017; AAF02489.1; -
CC EMBL; AF224724; AAF34712.1; -
CC HSSP; P00763; IDPO.
CC MEROPS; S01_228; -
DR MGD; MGI:1859281; Hgfac.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR000562; FN Type II.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; ENTYRE11.

DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR ProDom; PD000995; FN Type_II; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYPSIN; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Hydrolase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
KW EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 29
FT PROPEP 30 369
FT CHAIN 370 405
FT CHAIN 406 653
FT CHAIN 653
FT DOMAIN 105 145
FT DOMAIN 157 195
FT DOMAIN 197 237
FT DOMAIN 238 276
FT DOMAIN 283 364
FT DOMAIN 406 653
FT ACT SITE 445 445
FT ACT SITE 496 495
FT ACT SITE 595 596
FT DISULFID 105 130
FT DISULFID 119 145
FT DISULFID 161 172
FT DISULFID 166 183
FT DISULFID 185 194
FT DISULFID 199 227
FT DISULFID 225 234
FT DISULFID 242 253
FT DISULFID 247 264
FT DISULFID 266 275
FT DISULFID 283 364
FT DISULFID 304 346
FT DISULFID 335 359
FT DISULFID 392 519
FT DISULFID 430 446
FT DISULFID 438 508
FT DISULFID 533 602
FT DISULFID 552 581
FT DISULFID 565 620
FT CARBOHYD 39 39
FT CARBOHYD 47 47
FT CARBOHYD 63 63
FT CARBOHYD 287 287
FT CARBOHYD 466 466
FT CARBOHYD 544 544
FT CONFLICT 164 164
SQ SEQUENCE 653 AA; 70567 MW; 88B4B205DF7FDC CRC64;
Query Match 10.8%; Score 150; DB 1; Length 653;
Best Local Similarity 32.6%; Pred. No. 3; Ie-05;
Matches 46; Conservative 17; Mismatches 46; Indels 32; Gaps 7;
QY 25 CPWMDGHLRYEDQTSAPAGLRLNW-----LDAOSGLASAPVSGAGNHSYCRNPDED 76
DB 283 CFLGNGTEYRGVASTAAGLSCLAMNSDLYOEJHVDG-VAAAVLLGLGPHAYCRPDDX 341
QY 77 PRGPWCYVSGEAGVEKRPCCDLRCPEPTTSQALPAFTTEIQEASBEGGADVEOVFAPANA 136
DB 342 ER-FWCYVVKXNALSWE-----YCRLTACESLARVHSQTEP-----ILA---A 380

OY 137 LPARSEAAVQPVIGISQRR 157
DB 381 LP--ESAPAVRPTGKRRKR 399

RESULT 5
TPA_MOUSE STANDARD; PRT; 559 AA.
ID P1234;
AC 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (CPA)
DE (c-PA) (c-plasminogen activator).
GN PLAT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8087303; PubMed=2826484;
RA Rickles R.J., Darrow A.L., Strickland S.;
RT "Molecular cloning of complementary DNA to mouse tissue plasminogen
RT activator mRNA and its expression during F9 teratocarcinoma cell
RT differentiation.";
RL J. Biol. Chem. 263:1563-1569(1988).
CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOMEN PLASMINOGEN
CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC MANY OTHER PHYSIOLOGICAL EVENTS.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC BOND.
CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -1- PFM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE 1 DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC
CC EMBL: J035520; AAA40470.1; --
CC PIR: A29941; A29941.
DR HSSP: P00750; 1A5H.
DR MGD: MGI:97610; Plac.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000083; Fibrinchnl.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00039; fnl_1.
DR Pfam: PF00051; kringle_2.
DR Pfam: PF00089; trypsin_1.
DR PRINTS: PR00722; CHYMOTRYPsin.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle_2.

DR SMART: SM00181; EGF_1.
DR SMART: SM00058; FNI_2.
DR SMART: SM00130; KR_2.
DR SMART: SM00020; Tryp_Spc_1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS01253; FIBRONECTIN_1; 1.
DR PROSITE: PS00021; KRINGLE_1; 2.
DR PROSITE: PS00070; KRINGLE_2; 2.
DR PROSITE: PS00140; TRYPsin_DOM; 1.
DR PROSITE: PS00134; TRYPsin_HIS; 1.
DR PROSITE: PS00135; TRYPsin_SER; 1.
KW plasminogen activation; Hydrolyse; Serine protease; Glycoprotein;
KM plasmin; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 17
FT PROPEP 18 29
FT CHAIN 30 559
FT CHAIN 30 308
FT CHAIN 309 559
FT DOMAIN 36 78
FT DOMAIN 79 117
FT DOMAIN 124 205
FT DOMAIN 213 294
FT DOMAIN 309 559
FT ACT_SITE 355 355
FT ACT_SITE 404 404
FT ACT_SITE 510 510
FT DISULFID 38 68
FT DISULFID 66 75
FT DISULFID 83 94
FT DISULFID 88 105
FT DISULFID 107 116
FT DISULFID 124 205
FT DISULFID 145 187
FT DISULFID 176 200
FT DISULFID 213 294
FT DISULFID 234 276
FT DISULFID 265 289
FT DISULFID 297 428
FT DISULFID 340 356
FT DISULFID 348 417
FT DISULFID 442 516
FT DISULFID 474 490
FT DISULFID 506 534
FT CARBOHYD 149 149
FT CARBOHYD 481 481
SQ SEQUENCE 559 AA; 63110 MW; 4ACE57DC6A282A5 CRC64.

Query Match 10.6%; Score 147.5; DB 1; Length 559;
Best local similarity 37.0%; Pred. No. 4,3e-05;
Matches 34; Conservative 11; Mismatches 34; Indels 13; Gaps 4;

OY 25 CFMDNGHYREDQSPAPGLRCUNWLDAGSGLASAPVS-----GAGNHSYCRNDE 75
DB 124 CFEEGGLTRGTMTASGAECLNW--NSSVSLKRYNRRPAKLGSLGNHNYCRNDR 181

OY 76 DRRGFWCVSGEAGVPEKRCEDLRCPETTSQ 107
DB 182 DLK-FWCVGF-KAGKYTFTEFCSTPACPKKSE 211

RESULT 6
HGFA_HUMAN STANDARD; PRT; 655 AA.
ID Q04756; Q14726;
AC 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGFA)
DE activator (HGFA).
GN HGFAc.

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC TISSUE=Liver, and Serum.
RX MEDLINE=93252878; Pubmed=7683665;
RA Miyazawa K., Shimomura T., Kitamura A., Kondo J., Morimoto Y.,
RA Kitamura N.;
RT "Molecular cloning and sequence analysis of the cDNA for a human
RT serine protease responsible for activation of hepatocyte growth
RT factor: structural similarity of the protease precursor to blood
RT coagulation factor XII."
RT J. Biol. Chem. 268:10024-10028(1993).
RN (2)
RF SEQUENCE OF 40-655 FROM N.A.
RA Zhao S., Odeh C.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY
CC -1- CONVERTING IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM.
CC -1- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A
CC DISULFIDE BOND.
CC -1- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN
CC PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM.
CC -1- TISSUE SPECIFICITY: LIVER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 IS THE INITIATOR.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isdb-sib.ch/announce/
CC or send an email to license@sib-sb.ch).
-----
CC EMBL: D14012; BAO3113.1; -
CC EMBL: Z69923; CA93803.1; -
CC F1R; A46688; A46688.
CC HSSP; P00763; IDPO.
CC MEROPS; S01.228; -
CC Genew; HGNC:4894; HGFAc.
CC MIM; 604552; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF-2.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR000083; FibrinctnI.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; ENTPEPIT.
DR ProDom; PD000395; Kringle; 1.
DR ProDom; PD000995; FN_Type_II; 1.
DR SMART; SM00181; EGF_2; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYF_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.

```

```

DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 30
FT PROPEP 31 372
FT CHAIN 373 407
FT CHAIN 408 655
FT DOMAIN 108 148
FT DOMAIN 160 198
FT DOMAIN 200 240
FT DOMAIN 241 279
FT DOMAIN 286 367
FT ACT_SITE 447 447
FT ACT_SITE 497 497
FT ACT_SITE 598 598
FT ACT_SITE 108 133
FT DISULFID 122 148
FT DISULFID 164 175
FT DISULFID 188 197
FT DISULFID 202 230
FT DISULFID 228 237
FT DISULFID 245 256
FT DISULFID 250 267
FT DISULFID 269 278
FT DISULFID 286 367
FT DISULFID 307 349
FT DISULFID 338 362
FT DISULFID 394 521
FT DISULFID 442 448
FT DISULFID 440 510
FT DISULFID 535 604
FT DISULFID 567 583
FT DISULFID 594 622
FT CARBOHYD 48 48
FT CARBOHYD 230 290
FT CARBOHYD 468 468
FT CARBOHYD 492 492
FT CARBOHYD 546 546
FT CONFLICT 644 644
SQ SEQUENCE 655 AA; 70681 MW; 2CF72F1B862ED7 CRC64;
Query Match 10.5%; Score 146.5; DB 1; Length 655;
Best local Similarity 36.9%; Pred.No.6.2e-05;
Matches 41; Conservative 10; Mismatches 41; Indels 19; Gaps 6;
OY 25 CFWDNGLHREIDTSPPAGRLCLNW-----LDAOSGLASPVSGAGNHSYCRNPED 76
DB 286 CPLNGGTGYRGVASTASAGLSCLANWSDLYOELHVDV-VGAALALGLGPHAYCNPND 344
OY 77 PRGPWCYVSGEAGVP---EKRPCEDLRCPETSOALPATTICE-ASEG 122
DB 345 ER-PWCYVVKOSALSMEXCRLEACESL---TRVQLSPDLLATLPAPAPG 390
RESULT 7
ID PLIN HUMAN STANDARD; PRT; 810 AA.
AC P00747;
DT 21-JUN-1986 (Rel. 01, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Plasmogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
GN PG6.
OS Homo sapiens (Human).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90202879; PubMed=2318848;
RA Petersen T.E., Martzen M.R., Ichinose A., Davie E.W.;
RT "Characterization of the gene for human plasminogen, a key proenzyme
RL in the fibrinolytic system.";
RN J Biol. Chem. 265:6104-6111(1990).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97162490; PubMed=3030813;
RA Forstgen M., Raden B., Israelsen M., Larsen K., Heden L.-O.;
RT "Molecular cloning and characterization of a full-length cDNA clone
RL for human plasminogen.";
RN FEBS Lett. 213:254-260(1987).
[3]
RP SEQUENCE OF 20-810.
RA Sottirup-Jensen L., Petersen T.E., Magnusson S.;
RL Submitted (JUL-1977) to the PIR data bank.
[4]
RP SEQUENCE OF 292-810 FROM N.A.
RX MEDLINE=85023311; PubMed=6148961;
RA Malinowski D.P., Sadler J.E., Davie E.W.;
RT "Characterization of a complementary deoxyribonucleic acid coding for
RL human and bovine plasminogen.";
RN Biochemistry 23:4243-4250(1984).
[5]
RP SEQUENCE OF 20-100.
RX MEDLINE=75093329; PubMed=122932;
RA Wiman B., Wallen P.;
RT "Structural relationship between 'glutamic acid' and 'lysine' forms
RL of human plasminogen and their interaction with the NH₂-terminal
RT activation peptide as studied by affinity chromatography.";
RN Eur. J. Biochem. 50:489-494(1975).
[6]
RP SEQUENCE OF 95-580; 581-626; 657-700 AND 722-810.
RX Sottirup-Jensen L., Claess H., Zajdel M., Petersen T.E., Magnusson S.;
RL (In) Davidson J.F., Rowan R.W., Samama M.W., Desnoyers P.C. (eds.);
RT Progress in chemical fibrinolysis and thrombolysis, pp.3:191-209;
RN Raven Press, New York (1978).
[7]
RP SEQUENCE OF 483-604.
RX MEDLINE=76043692; PubMed=126863;
RA Wiman B., Wallen P.;
RT "Amino-acid sequence of the cyanogen-bromide fragment from human
RL plasminogen that forms the linkage between the plasmin chains.";
RN Eur. J. Biochem. 58:539-547(1975).
[8]
RP SEQUENCE OF 581-810.
RX MEDLINE=77225245; PubMed=142009;
RA Wiman B.;
RT "Primary structure of the B-chain of human plasmin.";
RN Eur. J. Biochem. 76:129-137(1977).
[9]
RP ACTIVE SITE.
RX MEDLINE=73149248; PubMed=4694729;
RA Robbins K.C., Bernabe P., Arrazola L., Summaria L.;
RT "The primary structure of human plasminogen. II. The histidine loop
RL of human plasmin: light (b) chain active center histidine sequence.";
RN J Biol. Chem. 248:1631-1633(1973).
[10]
RP ACTIVE SITE.
RX MEDLINE=69234739; PubMed=4240117;
RA Grotskopf W.R., Summaria L., Robbins K.C.;
RT "Studies of the active center of human plasmin. Partial amino acid
RL sequence of a peptide containing the active center serine residue.";
RN J Biol. Chem. 244:3590-3597(1969).
[11]
RP OMEGA-AMINOCARBOXYLIC ACID-BINDING SITES.
RX MEDLINE=82213905; PubMed=6919539;
RL Trexler M., Vail Z., Patsy L.;

RT "Structure of the omega-aminocarboxylic acid-binding sites of human
RT plasminogen. Arginine 70 and aspartic acid 56 are essential for
RT binding of ligand by kringle 4.";
RL J Biol. Chem. 257:7401-7406(1982).
[12]
RP FIBRIN AND OMEGA-AMINOCARBOXYLIC ACID BINDING SITES.
RX MEDLINE=85054794; PubMed=6094526;
RA Vail Z., Patsy L.;
RT "The fibrin-binding site of human plasminogen. Arginines 32 and 34
RL are essential for fibrin affinity of the kringle 1 domain.";
RN J Biol. Chem. 259:13690-13694(1984).
[13]
RP PHOSPHORYLATION SITE SER-597.
RX MEDLINE=97345939; PubMed=9201958;
RA Wang H., Proctor M., Bretchaue R.K., Castellino F.J.;
RT "Serine-578 is a major phosphorylation locus in human plasma
RL plasminogen.";
RN Biochemistry 36:8100-8106(1997).
[14]
RP CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=88185329; PubMed=3356193;
RA Marti T., Schaller U., Rickli E.E., Schmid K., Kamberling J.P.,
RA Gerwig G.J., van Halbeek H., Vliegenhart J.F.;
RT "The N- and O-linked carbohydrate chains of human, bovine and porcine
RL plasminogen. Species specificity in relation to alkylation and
RL fucosylation patterns.";
RN Eur. J. Biochem. 173:57-63(1988).
[15]
RP CARBOHYDRATE-LINKAGE SITE SER-268.
RX MEDLINE=97207306; PubMed=9054441;
RA Pitter-Sharp S.R., Stevens R.D., Andon N.L., Enghild J.J.,
RA Pizzo S.V.;
RT "Evidence for a novel O-linked sialylated triasaccharide on Ser-248 of
RL human plasminogen 2.";
RN J Biol. Chem. 272:7408-7411(1997).
[16]
RP CHARACTERIZATION OF ANGIOSTATIN AND PARTIAL SEQUENCE.
RX MEDLINE=95042728; PubMed=7525077;
RA O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,
RA Moses H., Lane W.S., Cao Y., Sage E.H., Folkman R.;
RT "Angiostatin: a novel angiogenesis inhibitor that mediates the
RL suppression of metastases by a Lewis lung carcinoma.";
RN Cell 79:315-328(1994).
[17]
RP CHARACTERIZATION OF ANGIOSTATIN.
RX MEDLINE=97238710; PubMed=9102221;
RA Sim B.K., O'Reilly M.S., Liang H., Fortier A.H., He W., Maden J.W.,
RA Lapevych R., Nacy C.A.;
RT "A recombinant human angiostatin protein inhibits experimental primary
RL and metastatic cancer.";
RN Cancer Res. 57:1329-1334(1997).
[18]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 374-461.
RX MEDLINE=92031502; PubMed=1657148;
RA Mulichak A.M., Tulinsky A., Ravichandran K.G.;
RT "Crystal and molecular structure of human plasminogen kringle 4
RL refined at 1.9-A resolution.";
RN Biochemistry 30:10576-10588(1991).
[19]
RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 374-461.
RX MEDLINE=92031503; PubMed=1657149;
RA Wu T.-P., Padmanabhan K., Tulinsky A., Mulichak A.M.;
RT "The refined structure of the epsilon-aminocaproic acid complex of
RL human plasminogen kringle 4.";
RN Biochemistry 30:10589-10594(1991).
[20]
RP X-RAY CRYSTALLOGRAPHY (1.67 ANGSTROMS) OF 376-454.
RX Stee B., Yamano A., Whitlow M., Teeter W.M.;
RT "Structure of human plasminogen kringle 4 at 1.68 Angstrom and 277 K.
RL A possible structural role of disordered residues.";
RN Acta Crystallogr. D 53:169-178(1997).
[21]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 102-181.

RX MEDLINE-96180681; PubMed=8611560;
RA Mathews I.I., Vanderhoff-Hanover P., Castellino F.J., Tulinsky A.;
RT "Crystall structures of the recombinant kringle 1 domain of human
RT plasminogen in complexes with the ligands epsilon-aminocaproic acid
RT and trans-4-(aminomethyl)cyclohexane-1-carboxylic Acid.";
RL Biochemistry 35:2567-2576(1996).
RP [122] X-RAY CRYSTALLOGRAPHY (1.66 ANGSTROMS) OF 480-563.
RX MEDLINE-96198034; PubMed=9521645;
RA Chang J., Mochalkin I., McCance S.G., Cheng B., Tulinsky A.,
RA Castellino F.J.;
RT "Structure and ligand binding determinants of the recombinant kringle
RT 5 domain of human plasminogen.";
RL Biochemistry 37:3256-3271(1998).
RP [123]
RX STRUCTURE BY NMR OF 96-184.
RX MEDLINE-94237157; PubMed=8161475;
RA Rejzante M.R., Llinas M.;
RT "1H-NMR assignments and secondary structure of human plasminogen
RT kringle 1.";
RL Eur. J. Biochem. 221:927-937(1994).
RP [124]
RX STRUCTURE BY NMR OF 96-184.
RX MEDLINE-94237158; PubMed=8161476;
RA Rejzante M.R., Llinas M.;
RT "Solution structure of the epsilon-aminohexanoic acid complex of
RT human plasminogen kringle 1.";
RL Eur. J. Biochem. 221:939-949(1994).
RN [125]
RP STRUCTURE BY NMR OF 183-354.
RX MEDLINE-96194156; PubMed=8652577;
RA Soehndel S., Hu C.-K., Marti D., Affolter M., Schaller J., Llinas M.,
RA Rickli E.B.;
RT "Recombinant gene expression and 1H NMR characteristics of the
RT kringle (2 + 3) supermodule: spectroscopic/functional individuality
RT of plasminogen kringle domains";
RL Biochemistry 35:2357-2364(1996).
RN [126]
RP STRUCTURE BY NMR OF 374-461.
RX MEDLINE-90219023; PubMed=2157850;
RA Atkinson R.A., Williams R.J.P.;
RT "Solution structure of the kringle 4 domain from human plasminogen by
RT 1H nuclear magnetic resonance spectroscopy and distance geometry.";
RL J. Mol. Biol. 212:541-552(1990).
RN [127]
RP VARIANTS PHE-374 AND THR-620.
Query Match 10.2%; Score 142; DB 1; Length 810;
Best Local Similarity 37.2%; Pred. No. 0.00019;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;
QY 25 CPWDNGHLVREDQTSPPAGRLCINMLDA-----QSGLASAPVSGAGNHSYCNPPEDPRG 79
DB 103 CKTGNGKNGKGTMSKTKNGJTCOKMSTSPHRFRSPATHPSGL-ENNYCHNPNDPOG 161
QY 80 PWCYSGEAGVPERK--PCEDLRCPRE 103
DB 162 PWCYTTD-----PEKRYDYCDILECEE 183
RESULT 8
PLMN RAT STANDARD; PRT; 169 AA.
AC 001177;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Plasminogen (BC 3.4.21.7) (Fragment).
GN PLG
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE-91250378; PubMed=1645711;
RA Kenalas J.J., Møller S.P.;
RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a
RT receptor site for plasminogen.";
RL J. Biol. Chem. 266:10825-10829(1991).
CC -1- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC A PROTHROMBOTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION
CC AND INFLAMMATORY; IN OVULATION IT WEAKENS THE WALLS OF THE
CC GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC LAMININ AND VON WILLEBRAND FACTOR.
CC -1- CATALYTIC ACTIVITY: preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -1- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC -1- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M62832; AAA41884.1; -
DR PIR: A40522; A40522.
DR HSRP: P00747; 1PMK.
DR MEMOS: S01.233; -
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser. protease_Try.
DR Pfam: PF00051; Kringle_2.
DR Pfam: PF00095; Kringle_2.
DR SMART: SM00130; KR; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS0070; KRINGLE_2; 2.
DR PROSITE: PS00240; TRYPSIN_DOM; PARTIAL.
DR PROSITE: PS00134; TRYPSIN_HIS; PARTIAL.
DR PROSITE: PS00135; TRYPSIN_SER; PARTIAL.
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Repeat.
FT DOMAIN 1 1
FT NON_TER 1 1
FT 1 10 KRINGLE 3 (BY SIMILARITY).
FT DOMAIN 34 112 KRINGLE 4 (BY SIMILARITY).
FT 139 >169 KRINGLE 5 (BY SIMILARITY).
FT DISULFID 34 112 BY SIMILARITY.
FT DISULFID 55 95 BY SIMILARITY.
FT DISULFID 83 107 BY SIMILARITY.
FT NON_TER 169 169
SQ SEQUENCE 169 AA; 18401 MW; 77A54214C49D010C CRC64;
Query Match 9.9%; Score 138; DB 1; Length 169;
Best Local Similarity 31.9%; Pred. No. 7.2e-05;
Matches 35; Conservative 13; Mismatches 48; Indels 14; Gaps 5;
QY 25 CPWDNGHLVREDQTSPPAGRLCINMLDA-----QSGLASAPVSGAGNHSYCNPPEDPRG 79
DB 34 CYGNGKSYRGRSTNTNGKKCSQSWMTSPHSHSTPAPNPDGL-ENNYCHNPNDPOG 92
QY 60 PWCYSGEAGVPERK--PCEDLRCPRETSQALPARTETIOASGSGAEE 127
DB 93 PWCYTTD-----PSYRMEYCNLAKCSFTGGV--ASALVYQVVSARCTSE 136


```

DR MEROPS; S01.231; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM0020; TRYP_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Plasmogen activation; Hydrolase; Serine protease; Glycoprotein;
KM Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20
FT CHAIN 1 442
FT CHAIN 21 188
FT CHAIN 190 442
FT DOMAIN 29 65
FT DOMAIN 72 153
FT DOMAIN 154 189
FT DOMAIN 190 442
FT CATH 152 152
FT CATH 33 41
FT DISULFID 35 53
FT DISULFID 55 64
FT DISULFID 179 210
FT DISULFID 220 236
FT DISULFID 228 289
FT DISULFID 324 393
FT DISULFID 356 372
FT DISULFID 383 411
FT ACT_SITE 235 235
FT ACT_SITE 286 286
FT ACT_SITE 387 387
FT ACT_SITE 241 241
FT CONFLICT 242 242
FT CONFLICT 288 288
SQ SEQUENCE 442 AA; 4916 MW; E832CFE501321EE CRC64;

Query Match 9.7%; Score 135.5; DB 1; Length 442;
Beet local similarity 36.9%; Pred. No. 0.00035;
Matches 38; Conservative 7; Mismatches 37; Indels 21; Gaps 5;

OY 25 CFWNGHLYREDQTSPPAGLRCLNWDAGSL---ASAPVS---GAGNHSYCRNEDDP 77
DB 72 CFEQGHSHYRGKANTNTGGRPLPMNSATVLTNTYHARPDALQIGKKNYCRNDP-NQ 130
OY 78 RGPWCYVS-----GEAGVP-----EKRPEDIRCHETTSQ 107
DB 133 RRPWCYVGVGLKQVQCVWPNCSGGESHRRPVPYDKGNFSPTE 173

RESULT 11
URTB_DESRO STANDARD; PRT; 431 AA.
ID URTB_DESRO
AC P98121.
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator beta precursor (EC 3.4.21.68) (DSBA
DE beta).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
NCBI_Taxid=9430;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boisdol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RP CHARACTERIZATION.
RX MEDLINE=9393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boisdol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Balduz B., Wilt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC EXCLUSIVELY HAEMOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
CC AGENT.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: MONOMER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@1sb-sib.ch).
CC -----
DR EMBL; M63969; AAA31594.1; -.
DR HSP; P98119; IASI.
DR MEROPS; S01.239; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; EGF; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM0020; TRYP_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KM Plasmogen activation; Hydrolase; Serine protease; Glycoprotein;
KM Kringle; EGF-like domain; Signal; Multigene family.
FT SIGNAL 1 36
FT CHAIN 37 431
FT DOMAIN 37 75
FT DOMAIN 82 163
FT DOMAIN 179 431
FT ACT_SITE 226 226
FT ACT_SITE 275 275
FT ACT_SITE 382 382
FT DISULFID 41 52
FT DISULFID 46 63
FT DISULFID 65 74
FT DISULFID 82 163
BY SIMILARITY.

```


FT CONFLICT 403 403 N -> K (IN REF. 2).
 FT CONFLICT 417 417 Y -> H (IN REF. 2).
 FT CONFLICT 435 435 M -> R (IN REF. 2).
 SQ SEQUENCE 477 AA; 53719 MW; 17486555COE5077C CRC64;

Query Match 9.6%; Score 134; DB 1; Length 477;
 Best Local Similarity 38.6%; Pred. No. 0.0005;
 Matches 27; Conservative 9; Mismatches 20; Indels 14; Gaps 3;

OY 25 CFWNDGHIYREDQTSFAPGLCLNWLDAOSGL-----ASAPVSGAGNHSICRPD 74
 DB 128 CYKDGQVTVRGTSSTSGAGQICNW--NSNLTFRRTNGRSDALTITGLGHNYCRMPD 184
 OY 75 EDRPGWCYV 84
 DB 185 NNSK-FWCYV 193

RESULT 13
 HGFL_MOUSE STANDARD; PRT; 716 AA.

AC P26928;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hepatocyte growth factor-like protein precursor (Macrophage stimulatory protein) (MSP).
 GN MST1 OR HGFL.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;

RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Liver;
 RX MEDLINE=92002017; PubMed=1832957;
 RA Fritznier Degen S.J., Stuart L.A., Han S., Jamison C.S.;
 RT "Characterization of the mouse cDNA and gene coding for a hepatocyte growth factor-like protein: expression during development."
 RL Biochem J 30:9781-9791(1991).
 CC -1- FUNCTION: PROBABLY HAS NO PROTEOLYTIC ACTIVITY, SINCE CRUCIAL AA CHARACTERISTIC OF SERINE PROTEASES CATALYTIC SITES ARE NOT CONSERVED.

CC -1- TISSUE SPECIFICITY: LIVER. LOWER LEVELS IN LUNG, PLACENTA AND ADRENAL.
 CC -1- DEVELOPMENTAL STAGE: IS EXPRESSED AT LOW LEVELS DURING GESTATION. JUST BEFORE BIRTH THE LEVEL INCREASES DRAMATICALLY AND REMAINS STABLE AFTERWARDS.
 CC -1- PTM: MAY BE CLEAVED AFTER AA 488, TO YIELD A TWO-CHAIN MOLECULE HELD TOGETHER BY DISULFIDE BONDS, OR CLEAVED INTO TWO SEPARATE POLYPEPTIDES.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; M74180; AAA50166.1; -;
 DR EMBL; M74181; AAA50167.1; -;
 DR HSSP; P00747; IREN.
 DR MEROPS; S01.975; -;
 DR MGD; MGI:96080; HGFL.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR003014; PAN.
 DR InterPro; IPR003609; Pan app.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00024; PAN; 1.

DR Pfam; PF00051; kringle; 4.
 DR Pfam; P00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR Prodom; PD000395; Kringle; 4.
 DR SMART; SM00130; KR; 4.
 DR SMART; SM00473; PAN AP; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PSS00021; KRINGLE_1; 4.
 DR PROSITE; PSS0070; KRINGLE_2; 4.
 DR PROSITE; PSS0240; TRYPsin_DOM; 1.
 DR Kringlie; Glycoprotein; Serine protease homolog; Repeat; Signal.
 KW Kringlie; Glycoprotein; Serine protease homolog; Repeat; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 716
 FT DOMAIN 19 109
 FT DOMAIN 110 186
 FT DOMAIN 191 268
 FT DOMAIN 292 370
 FT DOMAIN 379 457
 FT DOMAIN 489 716
 FT DISULFID 56 78
 FT DISULFID 60 66
 FT DISULFID 110 186
 FT DISULFID 131 169
 FT DISULFID 157 181
 FT DISULFID 191 268
 FT DISULFID 194 333
 FT DISULFID 212 251
 FT DISULFID 240 263
 FT DISULFID 292 370
 FT DISULFID 313 352
 FT DISULFID 341 364
 FT DISULFID 379 457
 FT DISULFID 400 440
 FT DISULFID 428 453
 FT DISULFID 477 593
 FT DISULFID 512 528
 FT DISULFID 607 672
 FT DISULFID 637 651
 FT DISULFID 662 690
 FT CARBOHYD 72 72
 FT CARBOHYD 173 173
 FT CARBOHYD 305 305
 FT CARBOHYD 620 620
 FT CONFLICT 19 19
 SQ SEQUENCE 716 AA; 80598 MW; BDC02EF85213ACC CRC64;

Query Match 9.6%; Score 134; DB 1; Length 716;
 Best Local Similarity 28.2%; Pred. No. 0.0008;
 Matches 42; Conservative 9; Mismatches 46; Indels 52; Gaps 7;

OY 2 LLAWVQAFVLSNMLAAVYSGS-----GCFWNDGHIYREDQTSFAPGLCLNWL 49
 DB 60 LLPTQ-----HSLHTLYHSSLSCHLFQKDYVTRCINDNGSVYRGYARFAGSLPQAW 134
 OY 50 ---LDAOSGLASAPVSGAGNHSICRPDPRPGWCYVS----- 85
 DB 135 SRFPNDHKYTPYRKGL-BENFCRPDDBPGWCYTTNSVRFQSGIKTCREAVCVL 193
 OY 86 -----GEAGVPEK-RPCE--DLRCPET 104
 DB 194 CNGEDYRGEVDVTVESGRECGRMQLQHPS 222

RESULT 14
 URUG_DESRO STANDARD; PRT; 394 AA.
 AC P49150;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Salivary plasminogen activator gamma precursor (EC 3.4.21.68) (DSPA gamma).

RT MEDLINE=84272706; PubMed=6589620;
RA Verde P., Scopelliti M.P., Galeffi P., di Nocera P., Blasi F.;
RT "Identification and primary sequence of an unspliced human urokinase
RT poly(A)+ RNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).
RN [8]
RP SEQUENCE OF 21-177.
RX MEDLINE=83055084; PubMed=6754569;
RA Gunler W.A., Steffens G.J., Oetting F., Kim S.-M.A., Frankus E.,
RT Flohe L.;
RT "The primary structure of high molecular mass urokinase from human
RT urine. The complete amino acid sequence of the A chain.";
RN [9]
RP SEQUENCE OF 156-176 AND 179-224.
RX MEDLINE=83003608; PubMed=6749491;
RA Schaller J., Nick H., Rickli B.E., Gillesen D., Lejgier W.,
RT Studer R.O.;
RT "Human low-molecular-weight urinary urokinase. Partial
RT characterization and preliminary sequence data of the two polypeptide
RT chains.";
RL Eur. J. Biochem. 125:251-257(1982).
RN [10]
RP SEQUENCE OF 158-410.
RX MEDLINE=83055099; PubMed=6754572;
RA Steffens G.J., Gunzler W.A., Oetting F., Frankus E., Flohe L.;
RT "The complete amino acid sequence of low molecular mass urokinase
RT from human urine.";
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=85000858; PubMed=8591045;
RA Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,
RT Dobson C.M., Stuart D.I., Jones E.Y.;
RT "The crystal structure of the catalytic domain of human
RT urokinase-type plasminogen activator.";
RL Structure 3:681-691(1995).
RN [12]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.
RX MEDLINE=20266327; PubMed=10805774;
RA Speil S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G.,
RT Bode W., Magdolen V., Huber R., Mordelet L.;
RT "(4-aminomethyl)phenylguanidine derivatives as nonpeptidic highly
RT selective inhibitors of human urokinase.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).
RN [13]
RP STRUCTURE BY NMR.
RX MEDLINE=89127526; PubMed=2536903;
RA Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.;
RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-
RT dimensional NMR.";
RL Nature 337:579-582(1989).
RN [14]
RP STRUCTURE BY NMR OF 67-155.
RX MEDLINE=93003110; PubMed=1327118;
RA Li X., Smith R.A.G., Dobson C.M.;
RT "Sequential 1H NMR assignments and secondary structure of the kringle
RT domain from urokinase.";
RL Biochemistry 31:9562-9571(1992).
RN [15]
RP STRUCTURE BY NMR OF 67-155.
RX MEDLINE=94149701; PubMed=8107091;
RA Li X., Bohman A.M., Llinas M., Smith R.A.G., Dobson C.M.;
RT "Solution structure of the kringle domain from urokinase-type
RT plasminogen activator.";
RL J. Mol. Biol. 235:1548-1559(1994).
RN [16]
RP VARIANT LEU-141.
RX MEDLINE=96186279; PubMed=8655231;
RA Yoshimoto M., Uchiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K.,
RT Sawasaki Y., Hanada K.;
RT "Characterization of single chain urokinase-type plasminogen
RT activator with a novel amino-acid substitution in the kringle

RT structure.";
RL Biochim. Biophys. Acta 1293:83-89(1996).
RN [17]
RP VARIANT LEU-141.
RX MEDLINE=97218551; PubMed=9065988;
RA Conne B., Berczy M., Belin D.;
RT "Detection of polymorphisms in the human urokinase-type plasminogen
RT activator gene.";
RL Thromb. Haemost. 77:434-435(1997).
RN [18]
RP ERRATUM.
RA Conne B., Berczy M., Belin D.;
RL Thromb. Haemost. 78:973-973(1997).
RN [19]
RP VARIANT LEU-141.
RX MEDLINE=97379720; PubMed=9194591;
RA Turkmen B., Schmitt W., Schmaljohr B., Trommler P., Hell W.,
RT Chutezberg S., Graef H., Magdolen V.;
RT "Mutational analysis of the gene encoding urokinase-type plasminogen
RT activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer.";
RL Electrophoresis 18:686-689(1997).
CC -1- FUNCTION: POTENT PLASMINOGEN ACTIVATOR AND IS CLINICALLY USED FOR
CC THERAPY OF THROMBOLYTIC DISORDERS.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN.
CC -1- PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used
CC in Pulmonary Embolism (PE) to initiate fibrinolysis.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; K02419; CA26268.1; -
DR EMBL; M15476; AAA61253.1; -
DR EMBL; D00244; BAA00175.1; -
DR EMBL; D1143; BAA01915.1; -
DR EMBL; X02760; CA26535.1; -
DR EMBL; AF377330; AAK53822.1; -
DR EMBL; BC033575; AAH13575.1; -
DR EMBL; K03266; AAC97138.1; -
DR EMBL; K02286; AAA61252.1; -
DR EMBL; A21571; CAA01559.1; -
DR EMBL; A18397; CAA01390.1; -
DR PIR; A00931; UKHU.
DR PIR; A32974; A32974.
DR PDB; 1KDU; 31-OCT-93.
DR PDB; 1LWU; 29-JAN-96.
DR PDB; 1URK; 08-MAY-95.
DR PDB; 1EJN; 17-MAY-00.
DR MEROPS; S01.231; -
DR GlycoSiteDB; P00749; -
DR Genew; HGNC:9052; PLAU.
DR MIM; 191840; -
DR InterPro; IPR003134; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; C1ypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.

DR SMART: SM00130: KR: 1.
DR SMART: SM00020: TRP_SPC: 1.

Query Match 9.5%; Score 132; DB 1; Length 431;
Best Local Similarity 32.7%; Pred No 0.00067;
Matches 33; Conservative 13; Mismatches 33; Indels 22; Gaps 5;

QY 25 CFMNGHLYREDOTSPARGLRWLDA-----QSGLASAPVSGAGNHSYCRNPDEDP 77
DB 70 CTBGNHFTYKAKASTUTWKRPLPMNSATYVLOQTYHAHRSDPLQGLGKANYCRNPD-NR 128
QY 78 RGPWCYVSGEAGVPEKRP-----CEDLRCPETTSOAL 109
DB 129 RRPWCYV--QVGL--KPLVQECNVHDCADGKKPSPEEL 164

Search completed: April 7, 2003, 06:48:34
Job time : 12.8979 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 06:57:52 ; Search time 55.4987 Seconds
(without alignments)
976,427 Million cell updates/sec

Title: US-10-057-951-2
Perfect score: 1393
Sequence: 1 MLAWVOAFVSNMLAEAY.....PVDPOBSGPTPLMGCAQTEGA 263

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP_Bacteria:*
2: SP_Bacteria:*
3: SP_fungi:*
4: SP_human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mhc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_virus:*
13: SP Vertebrate:*
14: SP_unclassified:*
15: SP_Virus:*
16: SP_Bacteriap:*
17: SP_Archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1393	100.0	263	4 Q96FE7	Q96FE7 homo sapien
2	1390	99.8	263	4 O00318	O00318 homo sapien
3	167	12.0	562	6 Q8SQ23	Q8SQ23 sus scrofa
4	154.5	11.1	516	4 Q9BU99	Q9BU99 homo sapien
5	150	10.8	653	11 Q8VCS4	Q8VCS4 mus musculu
6	147.5	10.6	559	11 Q8VCS4	Q8VCS4 mus musculu
7	142	10.2	810	4 Q15146	Q15146 mus musculu
8	139.5	10.0	385	5 Q25101	Q25101 herdmnia m
9	137	9.8	716	11 P70521	P70521 rattus norv
10	136.5	9.8	420	13 Q90504	Q90504 eptatretus
11	135	9.7	812	11 Q90504	Q90504 rattus norv
12	134.5	9.7	395	4 Q9BZM1	Q9BZM1 homo sapien
13	134.5	9.7	704	13 Q90865	Q90865 gallus galli
14	134	9.6	716	11 Q91XG8	Q91XG8 mus musculu
15	131.5	9.4	313	13 Q9PU78	Q9PU78 ctocodylus
16	131	9.4	154	4 Q96SE8	Q96SE8 homo sapien

17	131	9.4	608	13 Q9PTW7	Q9PTW7 struthio ca
18	130.5	9.4	717	13 P70006	P70006 xenopus lae
19	129	9.3	616	6 Q97507	Q97507 sus scrofa
20	128.5	9.2	806	6 Q18783	Q18783 macropus eu
21	127	9.1	157	6 Q9TVA8	Q9TVA8 bos taurus
22	126	9.0	716	13 Q91691	Q91691 xenopus lae
23	125.5	9.0	560	4 Q14520	Q14520 homo sapien
24	125	8.8	728	6 Q9BR09	Q9BR09 felis silve
25	121.5	8.7	399	4 Q96GL8	Q96GL8 homo sapien
26	121.5	8.7	420	4 Q9BTE9	Q9BTE9 homo sapien
27	119	8.5	334	6 Q46507	Q46507 papio hamed
28	119	8.5	710	13 Q91WJ5	Q91WJ5 mus musculu
29	117.5	8.4	812	11 Q91WJ5	Q91WJ5 mus musculu
30	114.5	8.2	113	4 Q9U1R7	Q9U1R7 homo sapien
31	114	8.2	113	4 Q9U1R5	Q9U1R5 homo sapien
32	114	8.2	202	13 Q90675	Q90675 gallus galli
33	114	8.2	452	13 Q90Y90	Q90Y90 xenopus lae
34	114	8.2	607	13 Q91001	Q91001 gallus galli
35	112.5	8.1	359	6 Q8MMR1	Q8MMR1 canis famli
36	112	8.0	567	4 Q13208	Q13208 caenorabdi
37	112	8.0	594	5 P91823	P91823 caenorabdi
38	112	8.0	1145	5 Q9BRL8	Q9BRL8 alysia cal
39	111.5	8.0	597	11 Q35727	Q35727 mus musculu
40	111	8.0	726	13 Q24488	Q24488 alysia cal
41	110.5	7.9	685	5 Q24488	Q24488 alysia cal
42	110.5	7.9	2358	16 Q91LV8	Q91LV8 streptomyce
43	108	7.8	215	13 Q42341	Q42341 gallus galli
44	106	7.6	378	13 Q90WPO	Q90WPO trachemys s
45	105.5	7.6	709	13 Q90ZNE	Q90ZNE brachydanto

ALIGNMENTS

RESULT 1

ID Q96FE7 PRELIMINARY: PRT; 263 AA.

AC Q96FE7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Unknown (protein for MGc:17330).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011049; AAH11049.1;
DR InterPro; IPR000001; Kinigle.
DR Pfam; PF00051; Kinigle; 1.
DR ProDom; PD000395; Kinigle; 1.
DR PROSITE; PS00021; KINIGLE_1; UNKNOWN_1.
DR PROSITE; PS50070; KINIGLE_2; 1.
SQ SEQUENCE 263 AA; 28234 MW; 197C3EBE888FA242 CRC64;

Query Match 100.0%; Score 1393; DB 4; Length 263;
Best Local Similarity 100.0%; Pred. No. 8.9e-118;
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MLAWVOAFVSNMLAEAYSGGCFWMDNGHLVREDQTSPPAGRLCLMWLDAOSGLASAP	60
DB	1	MLAWVOAFVSNMLAEAYSGGCFWMDNGHLVREDQTSPPAGRLCLMWLDAOSGLASAP	60
QY	61	VSGAGNHSYCRNDEDEPRGPGVCYSGAGVEKRPCEBDRCPETTSQALPFTTEIOEAS	120
DB	61	VSGAGNHSYCRNDEDEPRGPGVCYSGAGVEKRPCEBDRCPETTSQALPFTTEIOEAS	120
QY	121	EGPAGDEVQVPAPANALPKRSSTAAVQPVIGISQRYPMANSKEXKULGTLGVVIGITMMVI	180

```
DB 121 EGPADDEVQVAPANALPARSEAAAVDPVIGISQVRNMSKEKDLGLTGLVIGITMVI 180
QY 181 IIAIGAGIILIGSYKRGKDLKEQHDQKRCERMQRTTLPLSAFTNPCEIYDEKTVVHT 240
DB 181 IIAIGAGIILIGSYKRGKDLKEQHDQKRCERMQRTTLPLSAFTNPCEIYDEKTVVHT 240
QY 241 SQTPVDPQEGSTPLMGOAGTPGA 263
DB 241 SQTPVDPQEGSTPLMGOAGTPGA 263

RESULT 2
000318
ID 000318 PRELIMINARY; PRT; 263 AA.
AC 000318;
DT 01-JUN-1997 (TREMblrel. 04, Created)
DT 01-JUN-1997 (TREMblrel. 04, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE MUSC.DJ51SN1.2 protein.
CN MUSC.DJ51SN1.2.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Du Z., Scheet P., Harper M.;
RT "The sequence of H. sapiens PAC clone RP3-51SN1.";
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC002073; AAB54054.1; -.
DR HSSP: P00749; 1XDU.
DR InterPro: IPR000001; Kringle.
DR Pfam: PF00051; Kringle_1.
DR PRINTS: PR00010; KRINGLE; FALSE_NEG.
DR PRODOM: PD000395; Kringle; 1.
DR SMART: SM00130; KR; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
SQ SEQUENCE 263 AA; 28248 MW; 197C3EBE854A242 CRC64;

Query Match 99.8%; Score 1390; DB 4; Length 263;
Best Local Similarity 99.6%; Pred. No. 1.7e-117;
Matches 262; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAMVQAFVSNMLLAERYSGGCFNDGHLVREDQTSFPAQLRCINMLWDAQSGLASAP 60
DB 1 MLAMVQAFVSNMLLAERYSGGCFNDGHLVREDQTSFPAQLRCINMLWDAQSGLASAP 60
QY 61 VSGAGNHSYGRNDEDPGPMCVYSGRAGVPEKRPCEDELRCEPTTSQALPAFTTEIOAS 120
DB 61 VSGAGNHSYGRNDEDPGPMCVYSGRAGVPEKRPCEDELRCEPTTSQALPAFTTEIOAS 120
QY 121 EGPADDEVQVAPANALPARSEAAAVDPVIGISQVRNMSKEKDLGLTGLVIGITMVI 180
DB 121 EGPADDEVQVAPANALPARSEAAAVDPVIGISQVRNMSKEKDLGLTGLVIGITMVI 180
QY 181 IIAIGAGIILIGSYKRGKDLKEQHDQKRCERMQRTTLPLSAFTNPCEIYDEKTVVHT 240
DB 181 IIAIGAGIILIGSYKRGKDLKEQHDQKRCERMQRTTLPLSAFTNPCEIYDEKTVVHT 240
QY 241 SQTPVDPQEGSTPLMGOAGTPGA 263
DB 241 SQTPVDPQEGSTPLMGOAGTPGA 263

RESULT 3
08S023*
ID 08S023 PRELIMINARY; PRT; 562 AA.
AC 08S023; *
```

```
DT 01-JUN-2002 (TREMblrel. 21, Created)
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE T-plasminogen activator.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ENAMEL_ORGAN.
RA Ding Y., Xue J., Bartlett J.D.;
RT "T-plasminogen activator in tooth tissues.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF364605; AAM00297.1; -.
SQ SEQUENCE 562 AA; 63668 MW; F9E6B4C77CB101E8 CRC64;

Query Match 12.0%; Score 167; DB 6; Length 562;
Best Local Similarity 36.6%; Pred. No. 1.2e-06;
Matches 41; Conservative 12; Mismatches 33; Indels 26; Gaps 6;

QY 25 CFNDNGHLVREDQTSFPAQLRCINMLWDAQSGLASAPV-----GAGNHSYGRNPD 75
DB 127 CYEDQGITRYKGTWSTESGAECVNM--NTSGIASPYNGRPPAVKLGHNHYCRNPD 184
QY 76 DPGPMCVY-SGEAGVPEKRPCEDELRCEPTTSQALPAFTTEIOASEGPGAD 126
DB 185 DSR-PWCYIFKAEKSPD-----PC-----STPACTKEKECYTGKGLD 222

RESULT 4
09B099
ID 09B099 PRELIMINARY; PRT; 516 AA.
AC 09B099;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Similar to plasminogen activator, tissue.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SKIN;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: BC002795; AAH02795.1; -.
DR HSSP: P00750; 1ASH.
DR MEROPS: S01.242; -.
DR InterPro: IPR001214; Chymotrypsin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00051; Kringle; 2.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRODOM: PD00018; KRINGLE.
DR PRODOM: PD000395; Kringle; 2.
DR SMART: SM00181; EGF_1.
DR SMART: SM00001; EGF_like; 1.
DR SMART: SM00130; KR; 2.
DR SMART: SM00020; TRYPSIN; 1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00021; KRINGLE_1; UNKNOWN_2.
DR PROSITE: PS00070; KRINGLE_2; 2.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
```



```

DE Plasminogen precursor.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Liver;
RA Browne M.J., Chapman C.G., Dodd I., Carey J.E., Lawrence G.M.P.,
RA Mitchell D., Robinson J.H.;
RT "Expression of recombinant human plasminogen and aglycoplasminogen in
RT HeLa cells."; 0-0-0(1993).
RL Fibrinolysis.
CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, ALSO KNOWN AS THE
CC EMIL; M74220; AAA6451.1; -.
DR HSSP; P00747; 2PK4.
DR MEROPS; S01.233; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; PAN_app.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; Kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 5.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; TRY_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 5.
DR PROSITE; PS50070; KRINGLE_2; 5.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease; Signal.
FT SIGNAL 19 POTENTIAL.
FT CHAIN 20 810 PLASMINOGEN.
SQ SEQUENCE 810 AA; 90555 MW; B05C7DABD0020B3C CRC64;

Query Match 10.2%; Score 142; DB 4; Length 810;
Best Local Similarity 37.2%; Pred. No. 0.00035;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

QY 25 CFWDNGHLYREDQTSAPGLRCLMWLDA-----QSGLASAPVSGAGNHSYCRNPDEDPRG 79
DB 103 CRTGNGKNYRGITWSTKNGITCQKWSSTSPHREFRSPATHPEGL-BENYCRNPDDPG 161
QY 80 FWCYVSGEAGVPEKR--PCEDLRCE 103
DB 162 FWCYTTD---PEKRYDYCDILECEE 183

RESULT 8
Q25101 PRELIMINARY; PRT; 385 AA.
AC Q25101;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Serine protease.
GN HMERP1
OS Herdmania momus.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Pyrosidae; Herdmania.
OX NCBI_TaxID=7735;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CURVAIRA;
RA Arnold J.W., Kennett C., Javin M.F.;
RT "Transient expression of a novel serine protease in the ectoderm of

```

```

RT the ascidian Herdmania momus during development.";
RL Genes Evol. 206:455-463(1997).
CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; U63117; AAB6650.1; -.
DR HSSP; P00763; 1DPO.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; Kringle; 1.
DR Pfam; PF00024; PAN; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRY_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease.
SQ SEQUENCE 385 AA; 42935 MW; BFB1D05D5232E6A0 CRC64;

Query Match 10.0%; Score 139.5; DB 5; Length 385;
Best Local Similarity 20.2%; Pred. No. 0.00023;
Matches 70; Conservative 47; Mismatches 127; Indels 103; Gaps 14;

QY 2 ILAW--VOAFVSNMILABAYSGGCF-WDNGHLYREDQTSAPGLRCLMWLDAQS----- 54
DB 7 LVIWIIINGFVSN-----SECFDIENPESYGALSRIGGETQSW-DLQTPHNG 56
QY 55 -GLASAPVSGAGNHSYCRNPDEPRGWCYSGE-----AGVPEKRPEDDR 100
DB 57 KTSNENPENGGLANNYCNPNQDHWGWCYTNENFRMDYCDIPICSNPPVTLPSIE 116
QY 101 CEEY--SQALPAFTTEQESSECP---GADEVQ----- 129
DB 117 CKRTPELPDITGTYGLDKSRAKTNELHIVGTTVHSGIPIQVSLRAKELHFGGSI 176
QY 130 ----VFAPNALPANSAAVQPIYGISQVRMSKEKKDGLTGYVL----- 173
DB 177 LNRNWLTPAHCTRKQOPKYLALGDYDRIQDYSEKK---VGFLLFNHEKTPAT 232
QY 174 --GITMMVITIAIGAGILGYSYRKGDKLQHDQVC-----EREMQRTIL 218
DB 233 FENDITLWKDTSISATIFGQSVFPANRVPAAKSIIVSGMDPTKGTQDYKLVNQL 292
QY 219 PLSAFTNPCEIYDEKTV---VHTISQTPVDPQEGSTPLMGQACTP 261
DB 293 PYMSP--KLCKKLYSKVGAIVFETSLCAAYKKGKDSGQDSGSP 337

RESULT 9
P70521 PRELIMINARY; PRT; 716 AA.
AC P70521;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Macrophage stimulating protein precursor.
GN MSP
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RC MEDLINE=97011126; PubMed=8858136;
RA Ohshiro K., Iwama A., Matsuno K., Ezaki T., Sakamoto O., Hamaguchi I.,
RA Takasu N., Suda T.;
RT "Molecular cloning of Rat Macrophage-stimulating protein and its

```



```

RT involvement in the Male Reproductive System."
RL Biochem. Biophys. Res. Commun. 227:273-280(1996).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: X95096; CA64473.1; -.
DR HSSP: P00747; 1KRN.
DR MEROPS: S01.975; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000001; Kringles.
DR InterPro: IPR003609; Pan. app.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00051; Kringles; 4.
DR Pfam: PF00024; PAN; 1.
DR Pfam: PF00083; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR0018; KRINGLE.
DR PRODOM: PD000395; Kringles; 4.
DR SMART: SM00130; KR; 4.
DR SMART: SM00473; PAN_AP; 1.
DR PROSITE: PS00020; TRYPSIN; 1.
DR PROSITE: PS00021; KRINGLE_1; 4.
DR PROSITE: PS50070; KRINGLE_2; 4.
DR PROSITE: PS50240; TRYPSIN_DOM; 1.
DR HYDROLASE: Serine protease; Signal.
KT SIGNAL 1 31 POTENTIAL.
SQ SEQUENCE 716 AA; 80733 MW; 06B7DF3EF56D921F CRC64;

Query Match 9.8%; Score 137; DB 11; Length 716;
Best Local Similarity 27.3%; Pred. No. 0.00084;
Matches 48; Conservative 12; Mismatches 56; Indels 60; Gaps 8;

OY 2 ILAWQALVLSNMLAEVAGSG-----GCFWMDGHLVREDDTSAPGLRCLAW 49
DB 80 LRPWQ-----HSLRAQLHSSLCPLPKQDYATTCIMNGASVGTARAPDLPQGM 134
OY 50 ---LNAOSGLASAPVSGAGNHSYCRNPDEPRGWCY----- 65
DB 135 SRFPNDKHYTPYKNGL-BENFCNPDGDRGWCYTTIKRSRPOSGCISCREANCVW 193
OY 86 -----GEAGVPER-RPCE--DLRCPET-----TSQALPAFTTEIOASGSP 123
DB 194 CNGEDYRGEDVLTSGRECORDLPHSHFPEKFPDALKDNYCNPDASBP 249

RESULT 10
OQ90504 PRELIMINARY; PRT; 420 AA.
ID OQ90504;
AC OQ90504;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Thrombin.
OS Eptatretus scutell (Pacific hagfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Hyperotreti; Myxiniformes;
OC Myxiniidae; Eptatreinae; Eptatretus.
OC NCBI_TaxID=7765;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RC MEDLINE=92212913; PubMed=155783;
RA Banfield D.K., Macgillivray R.T.;
RT "Partial characterization of vertebrate prothrombin cDNAs:
RT amplification and sequence analysis of the B chain of thrombin from
RT nine different species."
RI Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
RI [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RC MEDLINE=94223694; PubMed=7513365;
RA Banfield D.K., Irwin D.M., Walz D.A., Macgillivray R.T.;
RT "Evolution of prothrombin: isolation and characterization of the cDNAs
RT encoding chicken and hagin prothrombin."

```

```

RL J. Mol. Evol. 38:177-187(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Banfield D.K.;
RT Submitted (DEC-1991) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: M81393; AAA21620.1; -.
DR HSSP: P00734; 1UWS.
DR MEROPS: S01.217; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000001; Kringles.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00051; Kringles; 1.
DR Pfam: PF00083; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR0018; KRINGLE.
DR PRODOM: PD000395; Kringles; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; TRYPSIN; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS50070; KRINGLE_2; 1.
DR PROSITE: PS50240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
DR HYDROLASE: Serine protease.
KV SEQUENCE 420 AA; 47888 MW; 64522AA21A57B67A CRC64;

Query Match 9.8%; Score 136.5; DB 13; Length 420;
Best Local Similarity 26.5%; Pred. No. 0.00047;
Matches 39; Conservative 16; Mismatches 77; Indels 15; Gaps 4;

OY 25 CFWMDGHLVREDDTSAPGLRCLMWDAOSGLASAPVSGAG-NHSYCRNPDEPRGWCY 83
DB 17 CVRREGDVRGLIMITWGKPLPMRGYSNPLPSOFTTAGLINSYCRNPDDSGWCY 76
OY 84 VSGEAVP-----EKRPEDLRCPETTSQALPAFTTEIOASBGRGADDEVQVAPANLPA 139
DB 77 TKVEGTVDVYQALNCE-----SGDIFVGTVDVYQALSGRSGAARLTLPFNKTFGN 129
OY 140 RSPAAAVPVGIGISGRVNRMSKEKDL 166
DB 130 GEECGKRPMPBLQK--NDRSDEL 153

RESULT 11
OQ90W03 PRELIMINARY; PRT; 812 AA.
ID OQ90W03;
AC OQ90W03;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Plasminogen protein precursor (EC 3.4.21.7).
OS PLASMINOGEN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RC TISSUE=LIVER;
RA Bangert K., Johnsen A.H., Thorsen S.;
RT "Rat plasminogen cDNA and gene structure."
RT Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.
RI [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RC MEDLINE=91250378; PubMed=1645711;
RA Kanalas U.O., Makret S.P.;
RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a
RT receptor site for plasminogen."
RI J. Biol. Chem. 266:10825-10829(1991).

```

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPsin FAMILY.
CC EMBL: AJ242649; CAB46014.1; -.
DR HSSP: P00747; 1PMK.
DR MEROPS: S01.233; -.
DR InterPro: IPR001114; Chymotrypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan_app.
DR InterPro: IPR001254; Ser_protease_Try.
DR InterPro: IPR001400; Somatostatin.
DR Pfam: PF00051; Kringle; 5.
DR Pfam: PF00024; PAN; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle; 5.
DR SMART: SM00130; KR; 4.
DR SMART: SM00473; PAN_AP; 1.
DR SMART: SM00020; TRYD_SPE; 1.
DR PROSITE: PS00021; KRINGLE_1; 5.
DR PROSITE: PS00070; KRINGLE_2; 5.
DR PROSITE: PS00338; SOMATOSTATIN_2; UNKNOWN_1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease; Signal.
FT SIGNAL 19
FT CHAIN 1
SQ SEQUENCE 812 AA; 90535 MW; 8C703C5110BEC9E CRC64;

Query Match 9.7%; Score 135; DB 11; Length 812;
Best Local Similarity 30.9%; Pred No 0.0015;
Matches 34; Conservative 14; Mismatches 48; Indels 14; Gaps 5.

QY 25 CFPMNDHLREQGTSPAPGLACLANLDA-----QSGLASAPVSGAGNHSYCNPPEDPRG 79
DB 376 CYGNGKSKRGSSSTTNTGKKCSQWSTPHSHSKTPANFEDAGL-EKNYCNPNPDORG 434
QY 80 PWCYVSGEANGPERK--PCEDLRCPETTSQALPAFTTEIQBASBQGADE 127
DB 435 PWCFTTD---PSYRWEXCNLKRCSSETGSGV--AESALVPQVPSAGTSE 478

RESULT 12
Q9BZW1 PRELIMINARY; PRT; 395 AA.
AC Q9BZW1;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Neocatal thrombolytic agent alpha-form (Fragment).
OS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
CC NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Dou D.;
RT "A brain-type plasminogen activator";
RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ database
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPsin FAMILY.
CC EMBL: AF260825; MAK11956.1; -.
DR HSSP: P00750; 1PK2.
DR MEROPS: S01.232; -.
DR InterPro: IPR001514; Chymotrypsin.
DR InterPro: IPR000083; Fibronin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00039; fnl; 1.
DR Pfam: PF00051; Kringle; 1.
DR Pfam: PF00089; trypsin; 1.

```

DR PRINTS; PR000172; CHYMOTRYPSIN.
DR PRINTS; PR00018; KIRINGLE.
DR ProDom; PD000395; Kringler; 1.
DR SMART; SM00058; FNI; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TYP SRC; 1.
DR PROSITE; PS001253; THERONECTIN_1; 1.
DR PROSITE; PS00072; KIRINGLE_2; 1.
DR PROSITE; PS00070; KIRINGLE_2; 1.
DR PROSITE; PS00240; TRYP SIN DOM; 1.
DR PROSITE; PS00134; TRYP SIN HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYP SIN SR; 1.
DR HydroLabe; Serine protease.
KW
FT
SQ
SEQUENCE 395 AA; 44323 MW; 3FBDA42F0B7C11C8 CRC64;

Query Match
Best Local Similarity 32.3%; Score 134.5; DB 4; Length 395;
Matches 32; Conservative 13; Mismatches 36; Indels 21; Gaps 4;

QY 4 AWQAFVLSNNL-----LAEYSGSGGCFMNDGHLREDOQTSPPAQRLCLNW----- 49
DB 18 SWLRPLVLSNRYEYECWNSGRACSEGNSDCYFENGSAVYRGTHTSLTESGASCLPPNMSML 77
OY 50 ---LDAQSGLASAPVSGAGNSYCRNFDEDPGPCVCY 84
DB 78 IGVYTAQN--PSAQLGLGKRNKCRNPDDAK-PNCHV 113

RESULT 13
090865
AC PRELIMINARY; PRT; 704 AA.
ID 090865;
AC 090865;
DT 01-NOV-1996 (TRENBERG; 01, Created)
DT 01-NOV-1996 (TRENBERG; 01, Last sequence update)
DT 01-MAR-2002 (TRENBERG; 20, Last annotation update)
GN Hepatocyte growth factor-like/macrophage stimulating protein.
OS HGFL/MSP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=96029010; PubMed=7554499;
RA Theys C., Sharpe M.J., Batley S.J., Stern C.D., Gherardi E.;
RT "Expression of HGF/SF, HGFL/MSP and c-met suggests new functions
RT during early chick development.";
RL Dev. Genet. 17:90-101(1995).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
EMBL; X84043; CAAS8862.1; -.
DR HSSP; P00747; ICEA.
DR MEROPS; S01.977; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringler.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan app.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00051; Kringler; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; CTRPSIN; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR0018; KIRINGLE.
DR ProDom; PD000395; Kringler; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN AP; 1.
DR SMART; SM00020; TYP SRC; 1.
DR PROSITE; PS00021; KIRINGLE_1; 4.
DR PROSITE; PS00070; KIRINGLE_2; 4.
DR PROSITE; PS00240; TRYP SIN DOM; 1.

```

KW Hydrolase; Serine protease.
SQ SEQUENCE 704 AA; 79341 MW; CAB0D8CC41367C37 CRC64;

Query Match 9.7%; Score 134.5; DB 13; Length 704;
Best Local Similarity 28.8%; Pred. No. 0.0014;
Matches 32; Conservative 11; Mismatches 45; Indels 23; Gaps 4;

QY 25 CFWDNGHLYREDQTSPPAPGLRCLNWLDAQSGLASAP-----VSGAGNHSYCRNPDED 76
DB 108 CIYANCTSYRGRDTERGRRCOHM-----QATTPHDFRFLPSLANGLEENYCRNPDED 161
QY 77 PRGPGCVV-----SGAGVPERKPCEDLRCPETTSQALPAFTTEIQEASE 121
DB 162 KRGPWCYTVDPNVRHOSCGI---KKCEDAVCMTCNGEDYRGFVDHTSGTE 209

RESULT 14

Q91XG8 PRELIMINARY; PRT; 716 AA.
AC Q91XG8;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hepatocyte growth factor-like.
GN HGFL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=LIVER;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010551; AAH10551.1; -.
DR MGD; MGI:96080; HGFL.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00051; kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR ProDom; PD000395; Kringle; 4.
DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_4.
DR PROSITE; PS50070; KRINGLE_2; 4.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 716 AA; 80693 MW; 12474C48A7D4B46D CRC64;

Query Match 9.6%; Score 134; DB 11; Length 716;
Best Local Similarity 28.2%; Pred. No. 0.0016;
Matches 42; Conservative 9; Mismatches 46; Indels 52; Gaps 7;

QY 2 LLAWVQAFLVSNMLLAAYSGS-----GCFWDNGHLYREDQTSPPAPGLRCLNW 49
DB 80 LLPWTQ-----HSLHTQLYHSSLCHLFOKKDYVRTCIMDNGVSRYGTAVARTAGGLPCQAW 134
QY 50 ---LDAQSGLASAPVSGAGNHSYCRNPDEDPRGWCYVS----- 85
DB 135 SRFPNDHKYTPTRKNGI--EENFCRNPDPGDPFGWCYTTNRSVRFOSCGIKTCREAVCVL 193
QY 86 -----GEAGVPERK-PPCE--DLRCPET 104
DB 194 CNGEDYRGFVDHTSGTEGRCQKMDLQHPHS 222

RESULT 15

Q9PU78 PRELIMINARY; PRT; 313 AA.
AC Q9PU78;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)

DE Hepatocyte growth factor-like protein (Fragment).
OS Crocodylus niloticus (Nile crocodile) (African crocodile).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylidae; Crocodylinae; Crocodylus.
OX NCBI_TaxId=8501;

RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=LIVER;
RC MEDLINE=20022983; PubMed=10555283;
RX Hughes S., Zelus D., Mouchiroud D.;
RA "Warm-blooded isochore structure in Nile crocodile and turtle.";
RT Mol. Biol. Evol. 16:1521-1527(1999).
RL EMBL; AJ011396; CAB56422.1; -.
DR HSSP; P00747; 1HPJ.
DR MEROPS; S01.977; -.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYPSIN; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
KW Hydrolase; Serine protease.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 313 AA; 34793 MW; 8E084704958B5AA2 CRC64;

Query Match 9.4%; Score 131.5; DB 13; Length 313;
Best Local Similarity 29.8%; Pred. No. 0.0092;
Matches 28; Conservative 11; Mismatches 42; Indels 13; Gaps 4;

QY 25 CFWDNGHLYREDQTSPPAPGLRCLNWLDAQSGLAS-APVSGAGNH---SYCRNPDEDPRG 80
DB 30 CYHNGELRYGHTSKTRKGVTCQKWSQSPRVQISPTTHPAHLDENYCRNPNDSHGP 89
QY 81 WCYVSGE-----AGVPERKPCEDLRCPETTSQA 108
DB 90 WCYTMDPRTPPDYCGI---KFCAGDKLPSVLENA 120

Search completed: April 7, 2003, 07:06:00
Job time : 59.4987 secs

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 06:41:55 ; Search time 41.3589 Seconds
(without alignments)
847.337 Million cell updates/sec

Title: US-10-057-951-2
Perfect score: 1393
Sequence: 1 MLAWQAFIVSNMLAEAY.....PVDPEQSTPLMGAGTGA 263

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1393	100.0	263	20	AAV05219	Kirngle1 protein s
2	1393	100.0	263	20	AAW87769	Human tissue plas
3	1393	100.0	263	22	AAE00300	Human tissue plas
4	1390	99.8	263	23	AAU86149	Human PRO264 poly
5	1384	99.4	263	21	AAU43337	Human ORF3001
6	1379	99.0	263	22	AAW93748	Human polypeptide
7	1167.5	83.8	286	20	AAV05220	Kirngle1 protein s
8	322	23.1	66	22	ABE31905	Peptide #311 enco
9	322	23.1	66	22	ABE3159	Protein #518 exp
10	322	23.1	66	22	AAW58537	Human brain expres

11	322	23.1	66	22	AAW1037	Human bone marrow
12	322	23.1	66	22	AAW18800	Peptide #5234 enco
13	322	23.1	66	22	AAW31314	Peptide #5351 enco
14	322	23.1	66	23	ABG40828	Human peptide enco
15	306	22.0	56	20	AA12615	Human 5' EST seque
16	305	21.9	55	20	AAV12397	Human 5' EST seque
17	196	14.1	39	19	AAW72641	Nervous glia cell
18	192	13.8	39	19	AAW72640	Nervous glia cell
19	160.5	11.5	527	13	AAW20317	t-PA analogue expr
20	160.5	11.5	527	13	AAW20320	t-PA analogue expr
21	160.5	11.5	527	13	AAW20218	t-PA analogue expr
22	160.5	11.5	527	13	AAW20219	t-PA analogue expr
23	160.5	11.5	527	8	AAW70059	Sequence of hybrid
24	160.5	11.5	650	8	AAW70060	Sequence of hybrid
25	160	11.5	351	20	AAW25407	Human tissue facto
26	159.5	11.5	472	10	AAW94416	Sequence of coding
27	159.5	11.5	527	13	AAW20223	t-PA analogue expr
28	159.5	11.5	527	13	AAW20222	t-PA analogue expr
29	159.5	11.5	527	13	AAW20221	t-PA analogue expr
30	157.5	11.3	527	19	AAW54154	t-PA mutant (N142S
31	157.5	11.3	527	19	AAW54157	t-PA mutant (N142S
32	157.5	11.3	562	10	AAW94379	Amino acid sequenc
33	157.5	11.3	562	10	AAW94380	Amino acid sequenc
34	156.5	11.2	439	16	AAW68851	Delta 2-89 tissue
35	156.5	11.2	483	16	AAW70869	Human tissue PA va
36	156.5	11.2	483	16	AAW70877	Human tissue PA va
37	156.5	11.2	483	16	AAW70878	Human tissue PA va
38	156.5	11.2	483	16	AAW70879	Human tissue PA va
39	156.5	11.2	483	16	AAW70880	Human tissue PA va
40	156.5	11.2	483	16	AAW70881	Human tissue PA va
41	156.5	11.2	483	16	AAW70882	Human tissue PA va
42	156.5	11.2	483	16	AAW70883	Human tissue PA va
43	156.5	11.2	483	16	AAW70884	Human tissue PA va
44	156.5	11.2	483	16	AAW70885	Human tissue PA va
45	156.5	11.2	483	16	AAW70886	Human tissue PA va

ALIGNMENTS

RESULT 1
AAV05219
ID AAV05219 standard; Protein; 263 AA.
XX
AC AAV05219;
XX
DT 17-JUN-1999 (first entry)
XX
DE Kirngle1 protein sequence.
XX
KW Kirngle1; cancer; inflammation; autoimmunity; rheumatoid arthritis; AIDS;
KW CNS inflammation; cerebellar degeneration; Alzheimer's disease; asthma;
KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
KW neurological abnormality; ischaemia reperfusion injury; ischaemic injury;
KW cardiovascular disease; kidney disease; liver disease; aplastic anaemia;
KW myocardial infarction; hypotension; hypertension; allergy; infection;
KW myelodysplastic syndrome; haematologic abnormality; diagnosis; therapy;
KW male pattern baldness.
XX
OS Homo sapiens.
XX
PN W09911788-A1.
XX
PD 11-MAR-1999.
XX
PF 02-SEP-1998; 98WC-US18270.
XX
PR 01-SEP-1998; 98US-0144889.
XX
PR 02-SEP-1997; 97US-0056032.
XX
PR (SMK) SMITHKLINE BEECHAM CORP.
XX
PI Albione EF, Kikly KK;

XX WPI: 1999-214707/18.
 DR N-PSDB; AAX28354.
 XX
 XX New kringle1 polypeptides and polynucleotides
 XX
 XX Claim 1; Page 31-32; 42pp; English.
 XX
 CC This sequence is a Kringle1 polypeptide of the invention.
 CC The kringle1 polypeptides (I) are used to screen for agonists and
 CC antagonists. Agonists are used to treat subjects in need of enhanced
 CC activity or expression of (I). Antagonists are used to treat subjects
 CC having need to inhibit the activity or expression of (I). The methods can
 CC be used to treat conditions such as cancer, inflammation, autoimmunity,
 CC allergy, asthma, rheumatoid arthritis, CNS inflammation, cerebellar
 CC degeneration, Alzheimer's disease, Parkinson's disease, multiple
 CC sclerosis, amyotrophic lateral sclerosis, head injury damage and other
 CC neurological abnormalities, ischemia reperfusion injury, cardiovascular
 CC disease, kidney disease, liver disease, ischemic injury, myocardial
 CC infarction, hypotension, hypertension, AIDS, myelodysplastic syndromes
 CC and other hematologic abnormalities, aplastic anaemia, male pattern
 CC baldness, and bacterial, fungal, protozoan and viral infections. The
 CC kringle1 polypeptides may also be used to generate antibodies.
 CC Determining the presence or absence of mutations in, and analysing for
 CC the presence or absence of expression of, kringle1 polynucleotides can be
 CC used to diagnose a disease or susceptibility to a disease related to
 CC expression or activity of kringle1 proteins. The polynucleotides may also
 CC be used for chromosome identification, and mapping.
 XX
 XX Sequence 263 AA;
 SQ
 Query Match 100.0%; Score 1393; DB 20; Length 263;
 Best Local Similarity 100.0%; Pred No. 9e-120;
 Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAMVQAFVSNMLAEYVSGGCFMNGHLYREDOETSAFRLCLMWLDNOSGLASAP 60
 DB 1 MLAMVQAFVSNMLAEYVSGGCFMNGHLYREDOETSAFRLCLMWLDNOSGLASAP 60
 QY 61 VSGAGNHSYCRNPDEDPKPGWCYVSGEAGVEKAPKPCDLRCPETTSQALPAFTTEIOEAS 120
 DB 61 VSGAGNHSYCRNPDEDPKPGWCYVSGEAGVEKAPKPCDLRCPETTSQALPAFTTEIOEAS 120
 QY 121 EGPAGDEVQVFAFANALPARSEAAVQPVISISQRYVMSKEKQDGLTGLVLTGITMVI 180
 DB 121 EGPAGDEVQVFAFANALPARSEAAVQPVISISQRYVMSKEKQDGLTGLVLTGITMVI 180
 QY 181 IIAIGAGIILGYSYKRGKDLKEOHOKYCEBEMORITLPLSAFTNPCEIVDEKTVVYHT 240
 DB 181 IIAIGAGIILGYSYKRGKDLKEOHOKYCEBEMORITLPLSAFTNPCEIVDEKTVVYHT 240
 QY 241 SQRTVDQEGESTPLMGQACTPCA 263
 DB 241 SQRTVDQEGESTPLMGQACTPCA 263

RESULT 2
 AAM87769
 ID AAM87769 standard; Protein; 263 AA.
 AC AAM87769;
 XX 29-MAR-1999 (first entry)
 XX
 DE Human tissue plasminogen activator-like protease t-PALP.
 XX
 KW Tissue plasminogen activator-like protease; t-PALP; human;
 KW circulatory system-related disorder; blood clotting; stroke;
 KW thrombosis; peripheral arterial occlusion; pulmonary embolism;
 KW myocardiothrombosis; diagnosis; therapy.
 XX
 XX Homo sapiens.
 XX

PH Key Location/Qualifiers
 FT Peptide 1..21
 FT /label= Sig_peptide
 FT Protein 22..263
 FT /label= Mac_protein
 FT Domain 25..84
 FT /note= "kringle domain"
 FT Domain 85..263
 FT /note= "protease domain"
 FT Peptide 22..31
 FT /note= "epitope-bearing region"
 FT Peptide 35..44
 FT /note= "epitope-bearing region"
 FT Peptide 71..81
 FT /note= "epitope-bearing region"
 FT Peptide 91..107
 FT /note= "epitope-bearing region"
 FT Peptide 119..128
 FT /note= "epitope-bearing region"
 FT Peptide 138..147
 FT /note= "epitope-bearing region"
 FT Peptide 155..167
 FT /note= "epitope-bearing region"
 FT Peptide 193..203
 FT /note= "epitope-bearing region"
 FT Peptide 206..215
 FT /note= "epitope-bearing region"
 FT Peptide 227..237
 FT /note= "epitope-bearing region"
 FT Peptide 243..252
 FT /note= "epitope-bearing region"
 FT
 PN MO9854199-A1.
 PD 03-DEC-1998.
 XX
 PF 27-MAY-1998; 98MO-US10728.
 XX
 PR 28-MAY-1997; 97US-0048000.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI
 PI Ebner R, Moore PA, Ruben SM;
 DR WPI: 1999-070207/06.
 DR N-PSDB; AAV98636.
 XX
 PT New tissue plasminogen activator-like protease - useful in the
 PT diagnosis and treatment of circulatory system-related disorders
 XX
 XX Claim 1; Page 56-57; 76pp; English.
 XX
 CC This is the amino acid sequence of tissue plasminogen activator-like
 CC protease (t-PALP), a novel member of the serine protease family
 CC that shares sequence homology to human tissue plasminogen activator
 CC (see AAM87770). The t-PALP sequence was deduced from a cDNA clone
 CC (see AAV98636) derived from activated monocytes. The 2.5 kb t-PALP
 CC message has also been detected in heart, brain, lung, placenta,
 CC liver, skeletal muscle, kidney, pancreas, spleen, thymus, prostate,
 CC testis, ovary, small intestine, colon and peripheral blood
 CC leukocytes. Isolated nucleic acids encoding amino acids -21 to
 CC 242, -30 to 242, 1-242, 4-63 (kringle domain) and 64-242 (protease
 CC domain) of t-PALP or encoding epitope-bearing portions of t-PALP,
 CC are also claimed, as are recombinant vectors, host cells, and
 CC methods for producing t-PALP polypeptides. t-PALP may be used to
 CC detect and treat disorders related to the circulatory system, and
 CC to identify agonists and antagonists of t-PALP activity. The
 CC homology between t-PALP and tPA indicates that t-PALP may be
 CC involved in the regulation of normal and abnormal clotting
 CC in e.g. stroke, deep-vein thrombosis, peripheral arterial
 CC occlusion, pulmonary embolism and myocardiothrombosis.
 CC
 XX
 XX Sequence 263 AA;

CC arterial occlusion, blood coagulation disorders, (auto)immune system
 CC disorders e.g. human immunodeficiency syndrome, rheumatoid arthritis,
 CC graft-versus-host disease, thyroiditis, insulin dependent diabetes and
 CC inflammatory eye disease, allergic reactions e.g. asthma, cardiovascular
 CC diseases e.g. heart disease, arrhythmia and myocardial ischemia,
 CC hyperproliferative disorders, cancers, hypertrophic scars and keloids,
 CC neurological diseases e.g. Creutzfeldt-Jakob syndrome, neurodegenerative
 CC disorders e.g. Alzheimer's disease and Parkinson's disease and infectious
 CC disease e.g. viral, bacterial and fungal infections. The t-PAP sequences
 CC are also useful for drug screening. The t-PAP nucleotides are useful as
 CC chromosome markers and are involved in gene therapy.

Sequence 263 AA:

Query Match 100.0%; Score 1393; DB 22; Length 263;
 Best Local Similarity 100.0%; Pred. No. 9e-120;
 Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAMVQAFIVSNMLIAEYSGGCFWMDNGLYREDQTSPPAGLRCLNMTDAQSGIASAP 60
 DB 1 MLAMVQAFIVSNMLIAEYSGGCFWMDNGLYREDQTSPPAGLRCLNMTDAQSGIASAP 60
 QY 61 VSGAGNSYCRNPDEDPGPMCVVSGAGVPEKRPCEDLRCPEETSQALPAFTTEIOEAS 120
 DB 61 VSGAGNSYCRNPDEDPGPMCVVSGAGVPEKRPCEDLRCPEETSQALPAFTTEIOEAS 120
 QY 121 EGPQADEVQVFAPANALPARSEAAAVQVIGISQVRNMSKEKDLGTGLGYVLGITMVI 180
 DB 121 EGPQADEVQVFAPANALPARSEAAAVQVIGISQVRNMSKEKDLGTGLGYVLGITMVI 180
 QY 181 IIAIGAGIILGYSYKRGKDLKEQHDQVRCEREMORTTLPLSAFNPFCIVDEKTVVHT 240
 DB 181 IIAIGAGIILGYSYKRGKDLKEQHDQVRCEREMORTTLPLSAFNPFCIVDEKTVVHT 240
 QY 241 SQTVPDPQEGSTPLMGQAGTPGA 263
 DB 241 SQTVPDPQEGSTPLMGQAGTPGA 263

RESULT 4

ID AAU86149 standard; Protein: 263 AA.

AC AAU86149;

DT 15-JUL-2002 (first entry)

DE Human PRO264 polypeptide.

KW Human; PRO; benign tumour; malignant tumour; lymphoid malignancy;
 KW leukaemia; neuronal disorder; stromal disorder; blastocytic disorder;
 KW inflammatory disorder; immune disorder; angiogenic disorder;
 KW cytoskeletal; neuroprotective.

OS Homo sapiens.

PN WO200153486-A1.

PD 26-JUL-2001.

PF 11-FEB-2000; 2000WO-US03565.

PR 08-MAR-1999; 99WO-US050328.

PR 11-MAR-1999; 99US-123972P.

PR 11-MAY-1999; 99US-133459P.

PR 02-JUN-1999; 99WO-US12252.

PR 22-JUN-1999; 99US-140650P.

PR 22-JUN-1999; 99US-140653P.

PR 20-JUL-1999; 99US-144758P.

PR 26-JUL-1999; 99US-145698P.

PR 28-JUL-1999; 99US-146222P.

PR 17-AUG-1999; 99US-149395P.

PR 31-AUG-1999; 99US-151689P.

PR 01-SEP-1999; 99WO-US20111.
 PR 15-SEP-1999; 99WO-US21090.
 PR 30-NOV-1999; 99WO-US28311.
 PR 01-DEC-1999; 99WO-US28301.
 PR 01-DEC-1999; 99WO-US28634.
 PR 05-JAN-2000; 2000WO-US00219.

PA (GENTH) GENENTECH INC.

PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
 PI Mazeres SA, Pan J, Plichi RM, Roy MA, Smith V, Stone DW;
 PI Matanabe CK, Wood WI;

DR WPI; 2002-205567/26.

DR N-PSDB; ABK40275.

XX Thirteen five nucleic acids encoding PRO polypeptides, useful for
 PT treating benign or malignant tumours, leukaemias and lymphoid
 PT malignancies, inflammatory, angiogenic and immunologic disorders -

PS Claim 61; Fig 44; 302pp; English.

CC The present invention relates to the isolation of novel human PRO
 CC polypeptides and the polynucleotide sequences encoding them. The
 CC PRO polypeptides, agonists, antagonists or anti-PRO antibodies are
 CC useful for treating benign or malignant tumours (e.g. renal, kidney,
 CC bladder, breast, etc), leukaemias and lymphoid malignancies, other
 CC disorders such as neuronal, glial, astrocytal, hypothalamic, glandular,
 CC macrophagal, stromal and blastocytic disorders, inflammatory, immune
 CC and angiogenic disorders. The polynucleotide sequences are also
 CC useful in gene therapy. AAU86128-AAU86162 represent the human PRO
 CC polypeptides of the invention.

Sequence 263 AA:

Query Match 99.8%; Score 1390; DB 23; Length 263;
 Best Local Similarity 99.8%; Pred. No. 1.7e-119;
 Matches 262; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAMVQAFIVSNMLIAEYSGGCFWMDNGLYREDQTSPPAGLRCLNMTDAQSGIASAP 60
 DB 1 MLAMVQAFIVSNMLIAEYSGGCFWMDNGLYREDQTSPPAGLRCLNMTDAQSGIASAP 60

QY 61 VSGAGNSYCRNPDEDPGPMCVVSGAGVPEKRPCEDLRCPEETSQALPAFTTEIOEAS 120
 DB 61 VSGAGNSYCRNPDEDPGPMCVVSGAGVPEKRPCEDLRCPEETSQALPAFTTEIOEAS 120

QY 121 EGPQADEVQVFAPANALPARSEAAAVQVIGISQVRNMSKEKDLGTGLGYVLGITMVI 180
 DB 121 EGPQADEVQVFAPANALPARSEAAAVQVIGISQVRNMSKEKDLGTGLGYVLGITMVI 180

QY 181 IIAIGAGIILGYSYKRGKDLKEQHDQVRCEREMORTTLPLSAFNPFCIVDEKTVVHT 240
 DB 181 IIAIGAGIILGYSYKRGKDLKEQHDQVRCEREMORTTLPLSAFNPFCIVDEKTVVHT 240

QY 241 SQTVPDPQEGSTPLMGQAGTPGA 263
 DB 241 SQTVPDPQEGSTPLMGQAGTPGA 263

RESULT 5

ID AAB43237 standard; Protein: 263 AA.

AC AAB43237;

DT 08-FEB-2001 (first entry)

DE Human ORFX ORF3001 polypeptide sequence SEQ ID NO:6002.

KW Human; open reading frame; ORFX; detection; cytoskeletal; hepatotropic;
 KW vulnerary; antiproliferative; antiparkinsonian; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;

KM immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KM hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KM antiviral; antibacterial; antifungal; antineumatic; antihypertensive;
 KM antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KM neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KM cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KM cholesterol ester storage; systemic lupus erythematosus; infection;
 KM severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KM allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KM bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KM thrombosis; contraceptive.
 OS Homo sapiens.
 XX
 XX MO200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000NO-US08621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-1) CURAGEN CORP.
 XX
 XX Shinketsu RA, Leach M;
 PI WPI; 2000-602362/57.
 DR N-PSDB; AAC77446.
 XX
 XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 XX Claim 11; Page 5181-5182; 5507pp; English.
 XX
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORF open reading frames 1 to 3161. The ORF
 CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;
 CC antiproliferative; antiparkinsonian; nootropic; neuroprotective;
 CC osteoplastic; anticonvulsant; antidiabetic; immunosuppressive;
 CC immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antineumatic;
 CC antihypertensive; antianaemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORF-associated disorder. The
 CC nucleic acids can be used to express ORF proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 SO Sequence 263 AA:
 Query Match 99.4%; Score 1384; DB 21; Length 263;
 Best Local Similarity 99.2%; Pred. No. 6e-119; Indels 0; Gaps 0;
 Matches 261; Conservative 1; Mismatches 1;
 QY 1 MLAAVQAFIVSNMLAEAYGGGCFWNGHLYREDOTSPAPGLCLAWLDAOSGLASAP 60
 DB 1 MLAAVQAFIVSNMLAEAYGGGCFWNGHLYREDOTSPAPGLCLAWLDAOSGLASAP 60
 QY 61 VSGAGNHSYCRNPDEPRGKWCYVSGEAGVPERKPCEDLRCPETTSQALPAFTTEIOEAS 120
 DB 61 VSGAGNHSYCRNPDEPRGKWCYVSGEAGVPERKPCEDLRCPETTSQALPAFTTEIOEAS 120

QY 121 EGGADDEVQVAPANALPARSEAAAVDPVIGISORVMSKEKKDLGTLGVLTITMMVI 180
 DB 121 EGGADDEVQVAPANALPARSEAAAVDPVIGISORVMSKEKKDLGTLGVLTITMMVI 180
 QY 181 IIAIGAGIILGYSYKRGKDLKEOHQKVCEREMORITLPSAFTNPCEIYDEKTVVHT 240
 DB 181 IIAIGAGIILGYSYKRGKDLKEOHQKVCEREMORITLPSAFTNPCEIYDEKTVVHT 240
 QY 241 SCPPVDPQEGSTPLMGAGCPGA 263
 DB 241 SCPPVDPQEGSTPLMGAGCPGA 263
 RESULT 6
 ID AAM93748
 XX AAM93748 standard; Protein; 263 AA.
 AC AAM93748;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polypeptide, SEQ ID NO: 3727.
 XX
 KM Human; full length cDNA; cDNA synthesis; oligo-capping.
 XX
 OS Homo sapiens.
 XX
 PN EP1130094-A2.
 XX
 PD 05-SEP-2001.
 XX
 PF 07-JUL-2000; 2000EP-0114089.
 XX
 PR 08-JUL-1999; 99JP-0194486.
 PR 11-JAN-2000; 2000JP-0118774.
 PR 02-MAY-2000; 2000JP-0183765.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Oka T, Nishikawa T, Inogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 DR WPI; 2001-524255/58.
 DR N-PSDB; AAK94700.
 XX
 PT 830 Primers useful for synthesizing full length cDNA clones and their
 PT use in genetic manipulation -
 XX
 XX Claim 8; SEQ ID NO 3727; 1380pp + sequence listing; English.
 XX
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesised by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is a polypeptide
 CC encoded by a full length human cDNA of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.
 XX
 SO Sequence 263 AA:
 Query Match 99.0%; Score 1379; DB 22; Length 263;
 Best Local Similarity 99.2%; Pred. No. 1.7e-116; Indels 0; Gaps 0;
 Matches 261; Conservative 0; Mismatches 2;
 QY 1 MLAAVQAFIVSNMLAEAYGGGCFWNGHLYREDOTSPAPGLCLAWLDAOSGLASAP 60
 DB 1 MLAAVQAFIVSNMLAEAYGGGCFWNGHLYREDOTSPAPGLCLAWLDAOSGLASAP 60

CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: the sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX

Sequence 66 AA;

Query Match 23.1%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 3,7e-22;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 TTSQLPAPFTTEIOASBGPADDEVGFAPANALPARSEAAAVPGTIGISQRYMNSKEK 163
DB 1 TTSQLPAPFTTEIOASBGPADDEVGFAPANALPARSEAAAVPGTIGISQRYMNSKEK 60
|||
QY 164 KDLGTL 169
|||
DB 61 KDLGTL 66

RESULT 9
ABB23159
ID ABB23159 standard; Protein; 66 AA.
XX
AC ABB23159;
XX

DT 23-JAN-2002 (first entry)
XX

DE Protein #158 encoded by probe for measuring heart cell gene expression.
XX

KM Human; gene expression; heart; microarray; vascular system;
KM cardiovascular disease; hypertension; cardiac arrhythmia;
KM congenital heart disease.
XX

OS Homo sapiens.
XX

PN WO200157274-A2.
XX

PD 09-AUG-2001.
XX

PF 30-JAN-2001; 2001WO-US00666.
XX

PR 04-FEB-2000; 2000US-0180312.
XX

PR 26-MAY-2000; 2000US-0207456.
XX

PR 30-JUN-2000; 2000US-0608408.
XX

PR 03-AUG-2000; 2000US-0632366.
XX

PR 21-SEP-2000; 2000US-0234687.
XX

PR 27-SEP-2000; 2000US-0236359.
XX

PR 04-OCT-2000; 2000GB-0024263.
XX

PA (MOLE-) MOLECULAR DYNAMICS INC.
XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX

DR WPI; 2001-488899/53.
XX

PT Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX

PS Claim 15; SEQ ID NO 24929; 530bp; English.
XX

CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the

CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

Sequence 66 AA;

Query Match 23.1%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 3,7e-22;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 TTSQLPAPFTTEIOASBGPADDEVGFAPANALPARSEAAAVPGTIGISQRYMNSKEK 163
DB 1 TTSQLPAPFTTEIOASBGPADDEVGFAPANALPARSEAAAVPGTIGISQRYMNSKEK 60
|||
QY 164 KDLGTL 169
|||
DB 61 KDLGTL 66

RESULT 10
AAM58537
ID AAM58537 standard; Protein; 66 AA.
XX
AC AAM58537;
XX

DT 05-NOV-2001 (first entry)
XX

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 30642.
XX

KM Human; brain expressed exon; gene expression analysis; probe;
KM microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KM epilepsy; cancer.
XX

OS Homo sapiens.
XX

PN WO200157275-A2.
XX

PD 09-AUG-2001.
XX

PF 30-JAN-2001; 2001WO-US00667.
XX

PR 04-FEB-2000; 2000US-0180312.
XX

PR 26-MAY-2000; 2000US-0207456.
XX

PR 30-JUN-2000; 2000US-0608408.
XX

PR 03-AUG-2000; 2000US-0632366.
XX

PR 21-SEP-2000; 2000US-0234687.
XX

PR 27-SEP-2000; 2000US-0236359.
XX

PR 04-OCT-2000; 2000GB-0024263.
XX

PA (MOLE-) MOLECULAR DYNAMICS INC.
XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX

DR WPI; 2001-483446/52.
XX

PT Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX

PS Example 4; SEQ ID NO: 30642; 650bp + sequence listing; English.
XX

CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancer. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX

Sequence 66 AA;
Query Match 23.1%; Score 322; DB 22; Length 66;

Best Local Similarity 100.0%; Pred. No. 3,7e-22;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 104 TTSQLPAFTTEIQEASBEGPDADEVQVAFAPANALPARSEAAAVQPIGTSQVRNNSKEK 163
DB 1 TTSQLPAFTTEIQEASBEGPDADEVQVAFAPANALPARSEAAAVQPIGTSQVRNNSKEK 60
OY 164 KDLGTL 169
DB 61 KDLGTL 66

RESULT 11

AAM71037
ID AAM71037 standard; Protein; 66 AA.

XX AAM71037;

DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed probe encoded protein SEQ ID NO: 31343.

KW Human; bone marrow expressed exon; gene expression analysis; probe;
microarray; cancer; leukemia; lymphoma; myeloma.

OS Homo sapiens.

PN WO200157276-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00668.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

PT Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 31343; 658bp + Sequence listing; English.

XX The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC bone marrow. They can be used to measure gene expression in bone marrow

CC samples, which may enable the improved diagnosis and treatment of cancers

CC such as lymphoma, leukemia and myeloma. The present sequence is a

CC protein encoded by one of the probes of the invention.

XX Sequence 66 AA;

XX

RESULT 12
AAM18800
ID AAM18800 standard; Protein; 66 AA.
XX AAM18800;
XX 12-OCT-2001 (first entry)
DT Peptide #5234 encoded by probe for measuring cervical gene expression.
DE Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer.
XX Homo sapiens.
OS WO200157278-A2.
PN 09-AUG-2001.
PD 30-JAN-2001; 2001WO-US00670.
PF 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488901/53.
DR Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX Claim 27; SEQ ID No 23626; 487bp; English.
XX The present invention relates to human single exon nucleic acid probes
CC (SENPs: see A110068-A1128459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at http://wipo.int/pub/published_pct_sequences.
XX Sequence 66 AA;
XX

Query Match 23.1%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 3,7e-22;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 104 TTSQLPAFTTEIQEASBEGPDADEVQVAFAPANALPARSEAAAVQPIGTSQVRNNSKEK 163
DB 1 TTSQLPAFTTEIQEASBEGPDADEVQVAFAPANALPARSEAAAVQPIGTSQVRNNSKEK 60
OY 164 KDLGTL 169
DB 61 KDLGTL 66

RESULT 13

AAM31314
ID AAM31314 standard; Protein; 66 AA.

XX AAM31314;

XX 17-OCT-2001 (first entry)
XX Peptide #5351 encoded by probe for measuring placental gene expression.
DE Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder.
XX Homo sapiens.
XX WO200157272-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US00663.
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488897/53.
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human placenta -
XX Claim 27; SEQ ID No 31583; 654pp; English.
XX The present invention relates to single exon nucleic acid probes (SENP:
XX see A1131315-A157546). The present sequence is a peptide encoded by one
XX such probe. The probes are useful for producing a microarray for
XX predicting, measuring and displaying gene expression in samples derived
XX from human placenta. The probes are useful for antenatal diagnosis of
XX human genetic disorders.
XX Sequence 66 AA;
SQ Query Match 23.1%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 3.7e-22;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 104 TTGQALPAFTTETQASSEGGADEVGFAPANALPARSEAAVQPVIGISORVRNNSKEK 163
DB 1 TTGQALPAFTTETQASSEGGADEVGFAPANALPARSEAAVQPVIGISORVRNNSKEK 60
QY 164 KDLGTL 169
DB 61 KDLGTL 66

RESULT 14
ABG40828
ID ABG40828 standard; Peptide; 66 AA.
XX
XX ABG40828;
XX
XX 19-AUG-2002 (first entry)
XX Human peptide encoded by genome-derived single exon probe SEQ ID 30493.
XX
XX Human; single exon probe; asthma; lung cancer; COPD; ILD;
XX Chronic obstructive pulmonary disease; interstitial lung disease;
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Kartagener syndrome;

KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX
XX Homo sapiens.
XX WO200186003-A2.
XX 15-NOV-2001.
XX 30-JAN-2001; 2001WO-US00665.
XX 04-FEB-2000; 2000US-180312P.
XX 26-MAY-2000; 2000US-207456P.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-234687P.
XX 27-SEP-2000; 2000US-236359P.
XX 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2002-114183/15.
XX Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples -
XX Claim 27; SEQ ID No 30493; 634pp; English.
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human lung comprising single exon nucleic acid probes having one of
XX 12614 nucleic acid sequences mentioned in the specification, or their
XX complements or the 12387 open reading frames derived from the 12614
XX probes. Also included are a microarray comprising the novel set of
XX probes; the novel set of probes which hybridise at high stringency to a
XX nucleic acid expressed in the human lung; measuring gene expression in a
XX sample derived from human lung, comprising (a) contacting the array with
XX a collection of detectably labeled nucleic acids derived from human lung
XX mRNA, and (b) measuring the label detectably bound to each probe of
XX the array; identifying exons in a eukaryotic genome, comprising
XX (a) algorithmically predicting at least one exon from genomic sequences
XX of the eukaryote; and (b) detecting specific hybridisation of detectably
XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
XX having a fragment identical to the predicted exon, the probe is included
XX in the above mentioned microarray; assigning exons to a single gene,
XX comprising (a) identifying exons from genomic sequence by the method
XX above and (b) measuring the expression of each of the exons in several
XX tissues and/or cell types using hybridisation to a single exon
XX microarray having a probe with the exon, where a common pattern of
XX expression of the exons in the tissues and/or cell types indicates that
XX the exons should be assigned to a single gene; a peptide comprising one
XX of 12011 sequences, mentioned in the specification, or encoded by the
XX probes/open reading frames (ORF). The probes are used for gene
XX expression analysis, and for identifying exons in a gene, particularly
XX using human lung derived mRNA and for the study of lung diseases
XX such as asthma, lung cancer, chronic obstructive pulmonary disease
XX (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
XX fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
XX Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
XX haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
XX pulmonary alveolar proteinosis, Kartagener syndrome, fibrocystic
XX pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
XX and hyaline membrane disease. The present sequence is a peptide/protein
XX encoded by a single exon probe of the invention.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic
XX format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 66 AA;

Query Match 23.1%; Score 322; DB 23; Length 66;
 Best Local Similarity 100.0%; Pred. No. 3,7e-22;
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 104 TTSGALPAFTTEIOEASGEGADEVOVFPANALPARSAAAVPVIGISQRYVMSKEK 163
 Db 1 TTSGALPAFTTEIOEASGEGADEVOVFPANALPARSAAAVPVIGISQRYVMSKEK 60

Oy 164 KDLCGL 169
 Db 61 KDLCGL 66

RESULT 15

AA12615
 ID AAY12615 standard; Protein; 56 AA.

XX AC AAY12615;

XX DT 22-JUN-1999 (first entry)

XX DE Human 5' EST secreted protein SEQ ID NO: 280 from WO 9906553.

XX KM Human; secreted protein; EST; expressed sequence tag; diagnosis;
 KM forensic; gene therapy; chromosome mapping; signal peptide;
 KM upstream regulatory sequence; cytokine activity; cell proliferation;
 KM differentiation; haematopoiesis regulation; tissue growth regulation;
 KM reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
 KM thrombolytic; antiinflammatory; tumour inhibition; antitumour.

XX OS Homo sapiens.

XX FN WO9906553-A2.

XX PD 11-FEB-1999.

XX PF 31-JUL-1998; 98WO-IB01237.

XX PR 01-AUG-1997; 97US-0905051.

XX PA (GEST) GENSET.

XX PI Duclert A, Dumas Milne Edwards J, Lacroix B;

XX DR WP1: 1999-153783/13.

XX DR N-PSDB; AAX41473.

XX PT New nucleic acids encoding human secreted proteins - obtained from
 PT cDNA libraries derived from umbilical cord, lymph ganglia,
 XX lymphocytes and placental tissue

XX PS Claim 34; Page 376; 411pp; English.

XX AAX41379 to AAX41526 represent 5' expressed sequence tags (ESTs) for
 CC human secreted proteins, and encode the proteins given in AAY12521 to
 CC AAY12668, respectively. The proteins given represent the signal peptide
 CC and an N-terminal fragment of a secreted protein. The nucleic acid
 CC sequences can be used for producing secreted human gene products. They
 CC can also be used to develop products for diagnosis and therapy. The
 CC proteins obtained may have cytokine activity, cell
 CC proliferation/differentiation activity, haematopoiesis regulating
 CC activity, tissue growth regulating activity, reproductive hormone
 CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
 CC thrombolytic activity, receptor/ligand activity, antiinflammatory
 CC activity, tumour inhibition activity or other activities. The products
 CC can be used in forensic, gene therapy and chromosome mapping procedures.
 CC The sequences can also be used for obtaining corresponding promoter
 CC sequences. The nucleic acids encoding the signal peptide can be used for
 CC directing extracellular secretion of a polypeptide or the insertion of a
 CC polypeptide into a membrane, or importing a polypeptide into a cell.

XX Sequence 56 AA;

Query Match 22.0%; Score 306; DB 20; Length 56;
 Best Local Similarity 98.2%; Pred. No. 8.8e-21;
 Matches 55; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MLAMVOAFIVSNMLAEAYSGGCFMDNGHLYREDQTSAPAGLRCLNWLDAQSG 56
 Db 1 MLAMVOAFIVSNMLAEAYSGGCFMDNGHLYREDQTSAPAGLRCLNWLDAQSG 56

Search completed: April 7, 2003, 06:59:47
 Job time : 43.3589 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 06:48:07 ; Search time 14.8468 Seconds

(without alignments)
1082.981 Million cell updates/sec

Title: US-10-057-951-2

Perfect score: 1393

Sequence: 1 MLAWVOAFLVSNMLAEAY.....PVDPOGSGPTLMGQGTGGA 263

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pap.*
2: /cgn2_6/ptodata/1/pubpaa/PCR_NEW_PUB.pap.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pap.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pap.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pap.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pap.*
7: /cgn2_6/ptodata/1/pubpaa/PCUTS_PUBCOMB.pap.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pap.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pap.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pap.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pap.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pap.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pap.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1393	100.0	263	9 US-10-102-704-2	Sequence 2, Appl1
2	1393	100.0	263	9 US-10-057-951-2	Sequence 2, Appl1
3	1393	100.0	263	10 US-09-084-491A-2	Sequence 2, Appl1
4	1393	23.1	527	9 US-09-864-457-18	Sequence 3, Appl1
5	154.5	11.1	527	9 US-09-867-455-18	Sequence 18, Appl1
6	154.5	11.1	527	9 US-09-867-455-18	Sequence 19, Appl1
7	154.5	11.1	527	9 US-09-867-455-18	Sequence 145, Appl1
8	154.5	11.1	527	9 US-09-867-455-18	Sequence 7, Appl1
9	142	10.2	160	9 US-09-335-325-35	Sequence 35, Appl1
10	142	10.2	160	9 US-09-335-325-35	Sequence 35, Appl1
11	142	10.2	160	9 US-09-335-325-35	Sequence 35, Appl1
12	142	10.2	160	9 US-09-335-325-35	Sequence 35, Appl1
13	142	10.2	250	9 US-10-131-241-30	Sequence 30, Appl1
14	142	10.2	250	9 US-10-131-241-30	Sequence 30, Appl1
15	142	10.2	250	9 US-10-131-241-30	Sequence 30, Appl1
16	142	10.2	250	9 US-10-131-241-30	Sequence 30, Appl1
17	142	10.2	339	9 US-09-335-325-3	Sequence 61, Appl1
18	142	10.2	339	9 US-10-131-241-3	Sequence 3, Appl1
19	142	10.2	339	10 US-09-761-120-3	Sequence 3, Appl1

20	142	10.2	352	9 US-09-335-325-40	Sequence 40, Appl1
21	142	10.2	352	9 US-10-131-241-40	Sequence 40, Appl1
22	142	10.2	352	10 US-09-761-120-40	Sequence 40, Appl1
23	142	10.2	368	10 US-09-761-120-42	Sequence 42, Appl1
24	142	10.2	378	9 US-09-335-325-42	Sequence 42, Appl1
25	142	10.2	378	9 US-10-131-241-42	Sequence 42, Appl1
26	142	10.2	378	9 US-09-873-676-1	Sequence 1, Appl1
27	142	10.2	458	10 US-09-946-893-4	Sequence 4, Appl1
28	142	10.2	569	10 US-09-946-893-5	Sequence 5, Appl1
29	142	10.2	571	10 US-09-946-893-8	Sequence 8, Appl1
30	142	10.2	576	10 US-09-946-893-6	Sequence 6, Appl1
31	142	10.2	791	9 US-09-967-186-1	Sequence 2, Appl1
32	142	10.2	810	9 US-09-946-893-2	Sequence 2, Appl1
33	138	9.9	79	9 US-09-335-325-8	Sequence 8, Appl1
34	138	9.9	79	9 US-10-131-241-8	Sequence 11, Appl1
35	138	9.9	79	9 US-09-967-455-11	Sequence 8, Appl1
36	138	9.9	79	10 US-09-753-064-2	Sequence 2, Appl1
37	135	9.7	377	9 US-09-987-455-8	Sequence 8, Appl1
38	133	9.5	326	9 US-10-057-951-3	Sequence 8, Appl1
39	133	9.5	354	9 US-09-987-457-10	Sequence 10, Appl1
40	133	9.5	354	9 US-09-987-455-11	Sequence 10, Appl1
41	133	9.5	372	9 US-10-102-704-3	Sequence 3, Appl1
42	132	9.5	372	10 US-09-084-491A-3	Sequence 3, Appl1
43	132	9.5	411	9 US-09-880-503-3	Sequence 3, Appl1
44	132	9.5	431	9 US-10-076-421-2	Sequence 2, Appl1
45	132	9.5	431	10 US-09-264-468B-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-10-102-704-2
; Sequence 2, Application US/10102704
; Patient No. US20020164768A1
; GENERAL INFORMATION:
; APPLICANT: Moore et al.
; TITLE OR INVENTION: Tissue Plasminogen Activator-Like Protein
; FILE REFERENCE: P378C1
; CURRENT APPLICATION NUMBER: US/10/102,704
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/084,491
; PRIOR FILING DATE: 1998-05-27
; PRIOR APPLICATION NUMBER: 60/048,000
; PRIOR FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-102-704-2

Query Match 100.0%; Score 1393; DB 9; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.6e-119; Gaps 0;
Matches 263; Conservative 0; Mismatches 0; Inlets 0; Gaps 0;
QY 1 MLAWVOAFLVSNMLAEAYGSGGCPMNGHLYREOTSPARGIACIMWLNQGLASAP 60
DB 1 MLAWVOAFLVSNMLAEAYGSGGCPMNGHLYREOTSPARGIACIMWLNQGLASAP 60
QY 61 VSGAONHSCNPNEDPRGPGWCYSGEAGVPERKPCEDLRCPETTSQALPAFTTEIOEAS 120
DB 61 VSGAONHSCNPNEDPRGPGWCYSGEAGVPERKPCEDLRCPETTSQALPAFTTEIOEAS 120
QY 121 EEPGADDEVQAPANNALPARSEAAAVQVIGISQVRVMSKEKKDGLTGLVYLITMWT 180
DB 121 EEPGADDEVQAPANNALPARSEAAAVQVIGISQVRVMSKEKKDGLTGLVYLITMWT 180
QY 181 IIAIAGIILLYSYKRGDKQKHQOKCEREMORITLPSAFNPTCEIIVDEKTVVHT 240
DB 181 IIAIAGIILLYSYKRGDKQKHQOKCEREMORITLPSAFNPTCEIIVDEKTVVHT 240

Mon Apr 7 07:36:46 2003

us-10-057-951-2.rapb

Page 2

QY 241 SQTVPDPOEGSTPLMGAGTGA 263
Db 241 SQTVPDPOEGSTPLMGAGTGA 263

RESULT 2

US-10-057-951-2
; Sequence 2, Application US/10057951
; Patent No. US20020177213A1
; GENERAL INFORMATION:
; APPLICANT: Moore et al.
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
; FILE REFERENCE: P378P1
; CURRENT APPLICATION NUMBER: US/10/057,951
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US 09/411,977
; PRIOR FILING DATE: 1999-10-04
; PRIOR APPLICATION NUMBER: US 09/084,491
; PRIOR FILING DATE: 1998-05-27
; PRIOR APPLICATION NUMBER: US 60/048,000
; PRIOR FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patentin Ver. 3.1
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-951-2

Query Match 100.0%; Score 1393; DB 9; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.6e-119; Indels 0; Gaps 0;
Matches 263; Conservative 0; Mismatches 0;

QY 1 MLAWVOAFVLSNMLAEAYSGGCFMNDGHLVREDQTSPPAGLRCLNMLDAGSGLASAP 60
Db 1 MLAWVOAFVLSNMLAEAYSGGCFMNDGHLVREDQTSPPAGLRCLNMLDAGSGLASAP 60
QY 61 VSGAGNHSYCRNDEDPGPMCVYSGEAGVPEKRPCEDLRCEFTTSQALPFTTEIOEAS 120
Db 61 VSGAGNHSYCRNDEDPGPMCVYSGEAGVPEKRPCEDLRCEFTTSQALPFTTEIOEAS 120
QY 121 EGPADDEVQVFAFANALPARSEAAAVOPVIGISQVRNMSKEKDLGTLGVIGITMVI 180
Db 121 EGPADDEVQVFAFANALPARSEAAAVOPVIGISQVRNMSKEKDLGTLGVIGITMVI 180
QY 181 IIAIGAGIIGYSYKRGKDLKEQHDQKVCEREMQRTITPLSAFTNPCEIYDEKTVVHT 240
Db 181 IIAIGAGIIGYSYKRGKDLKEQHDQKVCEREMQRTITPLSAFTNPCEIYDEKTVVHT 240
QY 241 SQTVPDPOEGSTPLMGAGTGA 263
Db 241 SQTVPDPOEGSTPLMGAGTGA 263

RESULT 3

US-09-084-491A-2
; Sequence 2, Application US/09084491A
; Patent No. US20020061576A1
; GENERAL INFORMATION:
; APPLICANT: MOORE, PAUL A.
; APPLICANT: RUBEN, STEVEN M.
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/084,491A
FILING DATE: 27-MAY-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: P378
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-084-491A-2

Query Match 100.0%; Score 1393; DB 10; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.6e-119; Indels 0; Gaps 0;
Matches 263; Conservative 0; Mismatches 0;

QY 1 MLAWVOAFVLSNMLAEAYSGGCFMNDGHLVREDQTSPPAGLRCLNMLDAGSGLASAP 60
Db 1 MLAWVOAFVLSNMLAEAYSGGCFMNDGHLVREDQTSPPAGLRCLNMLDAGSGLASAP 60
QY 61 VSGAGNHSYCRNDEDPGPMCVYSGEAGVPEKRPCEDLRCEFTTSQALPFTTEIOEAS 120
Db 61 VSGAGNHSYCRNDEDPGPMCVYSGEAGVPEKRPCEDLRCEFTTSQALPFTTEIOEAS 120
QY 121 EGPADDEVQVFAFANALPARSEAAAVOPVIGISQVRNMSKEKDLGTLGVIGITMVI 180
Db 121 EGPADDEVQVFAFANALPARSEAAAVOPVIGISQVRNMSKEKDLGTLGVIGITMVI 180
QY 181 IIAIGAGIIGYSYKRGKDLKEQHDQKVCEREMQRTITPLSAFTNPCEIYDEKTVVHT 240
Db 181 IIAIGAGIIGYSYKRGKDLKEQHDQKVCEREMQRTITPLSAFTNPCEIYDEKTVVHT 240
QY 241 SQTVPDPOEGSTPLMGAGTGA 263
Db 241 SQTVPDPOEGSTPLMGAGTGA 263

RESULT 4

US-09-864-761-38457
; Sequence 38457, Application US/09864761
; Patent No. US20020046763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30


```
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 38457
/ LENGTH: 66
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AC002073.1
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.5
/ OTHER INFORMATION: EXPRESSED IN HEEL, SIGNAL = 2.3
/ OTHER INFORMATION: EXPRESSED IN HEAT, SIGNAL = 1.9
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.2
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.6
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
/ OTHER INFORMATION: SWISSPROT HIT: P39881, EVALU6 3.40e+00
/ OTHER INFORMATION: EST_HUMAN HIT: BE299406.1, EVALU6 5.00e-30
/ US-09-864-761-38457

Query Match      23.1%; Score 322; DB 10; Length 66;
Best Local Similarity 100.0%; Pred. No. 1.6e-22;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 104 TTSSALPATTETIOGASRGPADDEVQVFAFANALPARSRAAVQPVIGISQRYMNSKEX 163
Db 1 TTSSALPATTETIOGASRGPADDEVQVFAFANALPARSRAAVQPVIGISQRYMNSKEX 60

Qy 164 KDLGTL 169
Db 61 KDLGTL 66

RESULT 5
US-09-987-457-18
/ Sequence 18, Application US/09987457
/ Publication No. US20030013150A1
/ GENERAL INFORMATION:
/ APPLICANT: Manosroi, Aranya
/ APPLICANT: Manosroi, Jiradej
/ APPLICANT: Tayapiwatana, Chatchai
/ APPLICANT: Goetz, Friedrich
/ APPLICANT: Werner, Rolf-Guenther
/ TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
/ FILE REFERENCE: 0652.2180001
/ CURRENT APPLICATION NUMBER: US/09/987,457
/ CURRENT FILING DATE: 2001-11-14
/ PRIOR APPLICATION NUMBER: 60/268,573
/ PRIOR FILING DATE: 2001-02-15
```

```
/ PRIOR APPLICATION NUMBER: GB 00 27 782.2
/ PRIOR FILING DATE: 2000-11-14
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: Patent In Ver. 2.1
/ SEQ ID NO 18
/ LENGTH: 527
/ TYPE: PRT
/ ORGANISM: Homo sapiens (tPA)
/ US-09-987-457-18

Query Match      11.1%; Score 154.5; DB 9; Length 527;
Best Local Similarity 39.6%; Pred. No. 4.6e-06;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

Qy 25 CFWDNGHLYREDQTSFAPGLRCLMWLDAOSGLASAPVS-----GAGNHSYCRNPDE 75
Db 92 CYEDQGISYRGTSVTAESGAECTNW--NSSALAQKPYSGRRPDAIRLGIQHNNYCRNPDR 149

Qy 76 DPRGPMCYVSGEAGVPEKRPCEDLRCPEPTTS 106
Db 150 DSK-PMCYVF-KAGKYSSEFCSTPACSEGNS 178

RESULT 7
US-09-974-298-145
/ Sequence 145, Application US/09974298
/ Patent No. US20020156263A1
/ GENERAL INFORMATION:
/ APPLICANT: Chen, Huei-Mei
/ APPLICANT: Chen, Huei-Mei
/ TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
/ FILE REFERENCE: PA-0037 P
/ CURRENT APPLICATION NUMBER: US/09/974,298
/ CURRENT FILING DATE: 2001-10-04

US-09-987-455-19
/ Sequence 19, Application US/09987455
/ Publication No. US20030049729A1
/ GENERAL INFORMATION:
/ APPLICANT: Aranya Manosroi
/ APPLICANT: Chatchai Tayapiwatana
/ APPLICANT: Friedrich Goetz
/ APPLICANT: Rolf-Guenther Werner
/ TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
/ FILE REFERENCE: 0652.2190001
/ CURRENT APPLICATION NUMBER: US/09/987,455
/ CURRENT FILING DATE: 2001-11-14
/ PRIOR APPLICATION NUMBER: 60/268,574
/ PRIOR FILING DATE: 2001-02-15
/ PRIOR APPLICATION NUMBER: GB 0027779.8
/ PRIOR FILING DATE: 2000-11-14
/ NUMBER OF SEQ ID NOS: 25
/ SOFTWARE: Patent In Ver. 2.1
/ SEQ ID NO 19
/ LENGTH: 527
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-09-987-455-19

Query Match      11.1%; Score 154.5; DB 9; Length 527;
Best Local Similarity 39.6%; Pred. No. 4.6e-06;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

Qy 25 CFWDNGHLYREDQTSFAPGLRCLMWLDAOSGLASAPVS-----GAGNHSYCRNPDE 75
Db 92 CYEDQGISYRGTSVTAESGAECTNW--NSSALAQKPYSGRRPDAIRLGIQHNNYCRNPDR 149

Qy 76 DPRGPMCYVSGEAGVPEKRPCEDLRCPEPTTS 106
Db 150 DSK-PMCYVF-KAGKYSSEFCSTPACSEGNS 178
```

PRIOR APPLICATION NUMBER: 60/238,331
PRIOR FILING DATE: 2000-05-10
NUMBER OF SEQ ID NOS: 194
SOFTWARE: PERL Program
SEQ ID NO 145
LENGTH: 562
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20020156263A1 1001470C01
US-09-974-298-145

Query Match 11.1%; Score 154.5; DB 9; Length 562;
Best Local Similarity 39.6%; Pred. No. 56-06; 36; Indels 13; Gaps 4;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 25 CFWMNGHLVREDQTSAPGLRCLMWLDAGSLASAPVS-----GAGNHSTCRNPDE 75
DB 127 CYEDQGISYRGWTWSTAESGAECTNW--NSSALAQKPYSGRRPDARLGLGNHNYCRNPDR 184
QY 76 DFRGPMWCYSGEAGVPEKRPCEDLRCPEPTTS 106
DB 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 213

RESULT 8
US-09-969-271-7
Sequence 7, Application US/09969271
Patent No. US20020098179A1
GENERAL INFORMATION:
APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
APPLICANT: Pfizer Limited (GB and EP (GB) only)
TITLE OF INVENTION: Pharmaceutical Combinations
FILE REFERENCE: PCS10951AEM
CURRENT APPLICATION NUMBER: US/09/969,271
CURRENT FILING DATE: 2001-10-01
PRIOR APPLICATION NUMBER: GB 0025473.0
PRIOR FILING DATE: 2000-10-17
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 562
TYPE: PRT
ORGANISM: Homo sapiens
US-09-969-271-7

Query Match 11.1%; Score 154.5; DB 10; Length 562;
Best Local Similarity 39.6%; Pred. No. 56-06; 36; Indels 13; Gaps 4;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 25 CFWMNGHLVREDQTSAPGLRCLMWLDAGSLASAPVS-----GAGNHSTCRNPDE 75
DB 127 CYEDQGISYRGWTWSTAESGAECTNW--NSSALAQKPYSGRRPDARLGLGNHNYCRNPDR 184
QY 76 DFRGPMWCYSGEAGVPEKRPCEDLRCPEPTTS 106
DB 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 213

RESULT 9
US-09-335-325-35
Sequence 35, Application US/09335325
Patent No. US20020164717A1
GENERAL INFORMATION:
APPLICANT: Folkman, M. Judah
O'Reilly, Michael
Gao, Yihai
Sim, B. Kim Lee
TITLE OF INVENTION: Angiostatin Fragments and Method of Use
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jones & Askev

STREET: 191 Peachtree Street, 37th Floor
CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/335,325
FILING DATE: 17-Jun-1999
CLASSIFICATION: <unknown>
PRIORITY INFORMATION:
APPLICATION NUMBER: US/08/612,788
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Warren, William L.
REGISTRATION NUMBER: 36,714
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 160 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
HYDROTHERMAL TYPE: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: K1-2
SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-335-325-35

Query Match 10.2%; Score 142; DB 9; Length 160;
Best Local Similarity 37.2%; Pred. No. 1.3e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

QY 25 CFWMNGHLVREDQTSAPGLRCLMWLDA-----QSLASAPVSGAGNHSTCRNPDEPRG 79
DB 1 CKTGNGKRYGTMKSKRNIGITCKMSSSTPHRRFSPATPSEGL-ENNYCRNPDPDQGS 59
QY 80 PWCYSGEAGVPEKRPCEDLRCPE 103
DB 60 PWCYTTD---PEKRYDYCDILECEE 81

RESULT 10
US-10-131-241-35
Sequence 35, Application US/10131241
Publication No. US20030012792A1
GENERAL INFORMATION:
APPLICANT: Holaday, John W.
APPLICANT: Fortlier, Anne H.
TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation Using Angiogenesis Inhibitors
FILE REFERENCE: 05213-0344 43170-271565
CURRENT APPLICATION NUMBER: US/10/131,241
CURRENT FILING DATE: 2002-07-22
PRIOR APPLICATION NUMBER: US 09/413,049
PRIOR FILING DATE: 1999-10-06
PRIOR APPLICATION NUMBER: US 09/316,802
PRIOR FILING DATE: 1999-05-21
PRIOR APPLICATION NUMBER: US 60/066,596
PRIOR FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 65

```

/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: U.S.
/ ZIP: 30303-1769
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/335,325
/ FILING DATE: 17-Jun-1999
/
/ CLASSIFICATION: <Unknown>
/
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/612,788
/ FILING DATE: <Unknown>
/
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warren, William L.
/ REGISTRATION NUMBER: 36,714
/ REFERENCE/DOCKET NUMBER: 05213-0126
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 404-818-3700
/ TELEFAX: 404-818-3799
/
/ INFORMATION FOR SEQ ID NO: 30:
/
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 250 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: <Unknown>
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHEICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: Homo sapiens
/ IMMEDIATE SOURCE:
/ CLONE: K1-3
/
/ SEQUENCE DESCRIPTION: SEQ ID NO: 30:
/
/ US-09-335-325-30
/
/
/ Query Match 10.2%; Score 142; DB 9; Length 250;
/ Best Local Similarity 37.2%; Pred. No. 2,4e-05;
/ Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;
/
/ QY 25 CFWDNGHLYREDQTSAPGLRCLMWLDA-----QSGLASAPVSGAGNHSYRNPDPEDPRG 79
/ Db 1 CKTGAGKNRYRTGMSYTXNGITCQKMSSTSPHRPRSPXTHSEGL-EEHYCRNPNDPQGS 59
/
/ QY 80 PWCYVSGEAGVEPKR--PCEDLRCE 103
/ Db 60 PWCYTTD---PEKRYDYCDILFCEE 81
/
/ RESULT 13
/ US-10-131-241-30
/ Sequence 30, Application US/10131241
/ Publication No. US20030012792A1
/
/ GENERAL INFORMATION:
/ APPLICANT: Holaday, John W.
/ TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolife
/ TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolife
/ FILE REFERENCE: 05213-0344 43170-271565
/ CURRENT APPLICATION NUMBER: US/10/131,241
/ CURRENT FILING DATE: 2002-07-22
/ PRIOR APPLICATION NUMBER: US 09/413,049
/ PRIOR FILING DATE: 1999-10-06
/ PRIOR APPLICATION NUMBER: US 09/316,802
/ PRIOR FILING DATE: 1999-05-21
/ PRIOR APPLICATION NUMBER: US 60/086,586
/ PRIOR FILING DATE: 1998-05-22
/ NUMBER OF SEQ ID NOS: 65
/
/ SOFTWARE: PatentIn version 3.1

```


GenCore version 5.1.4.B5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 06:37:51 ; Search time 192.301 Seconds

(without alignments)
881.767 Million cell updates/sec

Title: US-10-057-951-2

Perfect score: 1393

Sequence: MLAWVQAFVLSNMILAEAY.....FVDPQSGSTPLMGQAGTPGA 263

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 64473310 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1393	100.0	263	1	PCT-US00-27239-2
2	1393	100.0	263	1	PCT-US02-05301-164
3	1393	100.0	263	1	PCT-US98-10728-2
4	1393	100.0	263	14	US-09-084-491A-2
5	1393	100.0	263	15	US-09-144-889-2
6	1393	100.0	263	24	US-10-057-951-2

	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44																																																																																																																																																																																																																												
	1393	100.0	263	25	US-10-102-704-2	Sequence 2, Appl1	1393	100.0	263	27	US-60-048-000-2	Sequence 2, Appl1	1393	100.0	263	21	US-09-791-537-39562	Sequence 39562, A	1393	100.0	263	27	US-09-791-537-39562	Sequence 39562, A	1393	100.0	263	23	US-09-927-796-44	Sequence 44, Appl1	1393	100.0	263	26	US-10-210-951-44	Sequence 44, Appl1	1393	100.0	263	26	US-10-211-858-44	Sequence 44, Appl1	1393	100.0	263	26	US-10-211-858-44	Sequence 44, Appl1	1393	100.0	263	27	US-60-230-435-1580	Sequence 1580, App	1393	100.0	263	27	US-60-389-987-244	Sequence 244, App	1393	100.0	263	27	US-60-412-418-244	Sequence 244, App	1393	100.0	263	20	US-09-611-526-3727	Sequence 3727, App	1393	100.0	263	21	PCT-US01-14827-8858	Sequence 8858, App	1393	100.0	263	14	US-09-001-403-64	Sequence 8900, App	1393	100.0	263	19	PCT-US01-14827-8900	Sequence 8900, App	1393	100.0	263	25	US-60-207-315-471	Sequence 471, App	1393	100.0	263	15	US-09-144-889-4	Sequence 4, Appl1	1393	100.0	263	17	US-60-213-800-237	Sequence 237, App	1393	100.0	263	1	PCT-US02-05301-311	Sequence 311, App	1393	100.0	263	16	PCT-US02-05301-311	Sequence 311, App	1393	100.0	263	17	US-09-307-140-979	Sequence 979, App	1393	100.0	263	12	US-09-817-076-979	Sequence 979, App	1393	100.0	263	22	US-09-834-366-13839	Sequence 13839, A	1393	100.0	263	10	US-60-197-873-13839	Sequence 13839, A	1393	100.0	263	27	US-60-197-873-13839	Sequence 13839, A	1393	100.0	263	11	PCT-US01-14827-8897	Sequence 8897, App	1393	100.0	263	8	US-09-834-366-16200	Sequence 16200, A	1393	100.0	263	22	US-60-197-873-16200	Sequence 16200, A	1393	100.0	263	1	PCT-US01-00663-31583	Sequence 31583, A	1393	100.0	263	6	US-09-664-761-38457	Sequence 38457, A	1393	100.0	263	22	US-10-182-993-30642	Sequence 30642, A	1393	100.0	263	6	US-10-182-993-30642	Sequence 30642, A	1393	100.0	263	25	US-10-182-995-24929	Sequence 24929, A	1393	100.0	263	26	US-10-182-997-23626	Sequence 23626, A	1393	100.0	263	6	US-10-203-134-31343	Sequence 31343, A	1393	100.0	263	26	US-10-203-135-30493	Sequence 30493, A	1393	100.0	263	6	US-10-203-136-31400	Sequence 31400, A	1393	100.0	263	26	US-10-203-137-31583	Sequence 31583, A	1393	100.0	263	26	US-10-203-139-30540	Sequence 30540, A	1393	100.0	263	5	US-08-905-051-280	Sequence 280, App	1393	100.0	263	13	US-08-905-051-280	Sequence 280, App	1393	100.0	263	13	US-08-947-599C-2338	Sequence 2338, App	1393	100.0	263	5	US-08-905-135-428	Sequence 428, App

ALIGNMENTS

RESULT 1
PCT-US00-27239-2
; Sequence 2, Application PC/TUS0027239
; GENERAL INFORMATION:
; APPLICANT: HUMAN GENOME SCIENCES, INC.
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
; FILE REFERENCE: P378PCT2
; CURRENT APPLICATION NUMBER: PCT/US00/27239
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 1999-10-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US00-27239-2

QY	Query Match	Score	1393;	DB 1;	Length	263;
QY	1	MLAWVQAFVLSNMILAEAYGSGCFWMDNGHLYRBDQTSPPAGLRCLNWLDAQGLASAP	100.0%;	Best local similarity	100.0%;	Pred. No. 6.3e-129;
DB	1	MLAWVQAFVLSNMILAEAYGSGCFWMDNGHLYRBDQTSPPAGLRCLNWLDAQGLASAP	0;	Matches 263;	Conservative	0;
QY	61	VSQAGNHSYCRNPEDPRGWCYSGEAGVPERKPCDLCRCETTSQALPAFTTEIOEAS	100.0%;	Best local similarity	100.0%;	Pred. No. 6.3e-129;
DB	61	VSQAGNHSYCRNPEDPRGWCYSGEAGVPERKPCDLCRCETTSQALPAFTTEIOEAS	0;	Matches 263;	Conservative	0;

Db 61 VSGAGNHSYCRNPDEDPGRGWCYVSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEAS 120
QY 121 EGPADAEVQVFAPANALPARSEAAAVQPIVIGISQVRNNSKEKKDLGTLGVIGITMMVT 180
Db 121 EGPADAEVQVFAPANALPARSEAAAVQPIVIGISQVRNNSKEKKDLGTLGVIGITMMVT 180
QY 181 IIAIGAGIILIGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIYDEKTVVHT 240
Db 181 IIAIGAGIILIGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIYDEKTVVHT 240
QY 241 SQTVPDPOEGSTPLMGQAQTPGA 263
Db 241 SQTVPDPOEGSTPLMGQAQTPGA 263

RESULT 2

PCT-US02-05301-164
; Sequence 164, Application PC/TUS0205301
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: PS736PCT
; CURRENT APPLICATION NUMBER: PCT/US02/05301
; CURRENT FILING DATE: 2002-02-21
; PRIOR FILING DATE: 2001-07-12
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US 60/304,417
; PRIOR APPLICATION NUMBER: US 60/270,625
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 164
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-05301-164

Query Match 100.0%; Score 1393; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 6,3e-129;
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAWQAFIVSNMLLAAYVSGGCFWMDNGHLYREDQTSAPGLRCINMLDQSGLASAP 60
Db 1 MLAWQAFIVSNMLLAAYVSGGCFWMDNGHLYREDQTSAPGLRCINMLDQSGLASAP 60
QY 61 VSGAGNHSYCRNPDEDPGRGWCYVSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEAS 120
Db 61 VSGAGNHSYCRNPDEDPGRGWCYVSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEAS 120
QY 121 EGPADAEVQVFAPANALPARSEAAAVQPIVIGISQVRNNSKEKKDLGTLGVIGITMMVT 180
Db 121 EGPADAEVQVFAPANALPARSEAAAVQPIVIGISQVRNNSKEKKDLGTLGVIGITMMVT 180
QY 181 IIAIGAGIILIGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIYDEKTVVHT 240
Db 181 IIAIGAGIILIGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIYDEKTVVHT 240
QY 241 SQTVPDPOEGSTPLMGQAQTPGA 263
Db 241 SQTVPDPOEGSTPLMGQAQTPGA 263

RESULT 3

PCT-US98-10728-2
; Sequence 2, Application PC/TUS9810728
; GENERAL INFORMATION:
; APPLICANT: EBERNER, REINHARD
; APPLICANT: MOORE, PAUL
; APPLICANT: RUBEN, STEVEN
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE

; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/10728
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A.
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: P379PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8439
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US98-10728-2

Query Match 100.0%; Score 1393; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 6,3e-129;
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAWQAFIVSNMLLAAYVSGGCFWMDNGHLYREDQTSAPGLRCINMLDQSGLASAP 60
Db 1 MLAWQAFIVSNMLLAAYVSGGCFWMDNGHLYREDQTSAPGLRCINMLDQSGLASAP 60
QY 61 VSGAGNHSYCRNPDEDPGRGWCYVSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEAS 120
Db 61 VSGAGNHSYCRNPDEDPGRGWCYVSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEAS 120
QY 121 EGPADAEVQVFAPANALPARSEAAAVQPIVIGISQVRNNSKEKKDLGTLGVIGITMMVT 180
Db 121 EGPADAEVQVFAPANALPARSEAAAVQPIVIGISQVRNNSKEKKDLGTLGVIGITMMVT 180
QY 181 IIAIGAGIILIGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIYDEKTVVHT 240
Db 181 IIAIGAGIILIGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIYDEKTVVHT 240
QY 241 SQTVPDPOEGSTPLMGQAQTPGA 263
Db 241 SQTVPDPOEGSTPLMGQAQTPGA 263

RESULT 4

US-09-084-491A-2
; Sequence 2, Application US/09084491A
; GENERAL INFORMATION:
; APPLICANT: MOORE, PAUL A.
; APPLICANT: RUBEN, STEVEN M.
; APPLICANT: EBERNER, REINHARD
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/084,491A
; FILING DATE: 27-MAY-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A.
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF378
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-084-491A-2

Query Match 100.0%; Score 1393; DB 14; Length 263;
Best Local Similarity 100.0%; Pred. No. 6,3e-129;
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLTAWQAFIVSNMLAEAYSGGCFWMDNGHLVREDQTSAPGRLCLNMLDQSGLASAP 60
DB 1 MLTAWQAFIVSNMLAEAYSGGCFWMDNGHLVREDQTSAPGRLCLNMLDQSGLASAP 60
QY 61 VSGAGNHSYCRNPDEDPGRCVYSGEAGVPEKRPCEDLRCPTTSQALPAFTTEIOEAS 120
DB 61 VSGAGNHSYCRNPDEDPGRCVYSGEAGVPEKRPCEDLRCPTTSQALPAFTTEIOEAS 120
QY 121 EGPADDEVQFAPANALPARSEAAAVQPVIGISQVRNMSKEKKDLGTLGYLGITMMVI 180
DB 121 EGPADDEVQFAPANALPARSEAAAVQPVIGISQVRNMSKEKKDLGTLGYLGITMMVI 180
QY 181 IIAIGAGIILGYSYRGKDLKEQHDQVCEREMQRTPLPSAFTNPCEIYDEKTVVHT 240
DB 181 IIAIGAGIILGYSYRGKDLKEQHDQVCEREMQRTPLPSAFTNPCEIYDEKTVVHT 240
QY 241 SQTVPDQEGSTPLMGAGTPGA 263
DB 241 SQTVPDQEGSTPLMGAGTPGA 263

RESULT 5

US-09-144-889-2
; Sequence 2 Application US/09144889B
; GENERAL INFORMATION:
; APPLICANT: Earl F. Albone
; APPLICANT: Kristine K. Kikly
; TITLE OF INVENTION: KRINGLE1
; FILE REFERENCE: GH-70249
; CURRENT APPLICATION NUMBER: US/09/144, 889B
; CURRENT FILING DATE: 1998-09-01
; EARLIER APPLICATION NUMBER: 60/056, 032
; EARLIER FILING DATE: 1997-09-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-144-889-2

Query Match 100.0%; Score 1393; DB 15; Length 263;
Best Local Similarity 100.0%; Pred. No. 6,3e-129;
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLTAWQAFIVSNMLAEAYSGGCFWMDNGHLVREDQTSAPGRLCLNMLDQSGLASAP 60
DB 1 MLTAWQAFIVSNMLAEAYSGGCFWMDNGHLVREDQTSAPGRLCLNMLDQSGLASAP 60

QY 61 VSGAGNHSYCRNPDEDPGRCVYSGEAGVPEKRPCEDLRCPTTSQALPAFTTEIOEAS 120
DB 61 VSGAGNHSYCRNPDEDPGRCVYSGEAGVPEKRPCEDLRCPTTSQALPAFTTEIOEAS 120
QY 121 EGPADDEVQFAPANALPARSEAAAVQPVIGISQVRNMSKEKKDLGTLGYLGITMMVI 180
DB 121 EGPADDEVQFAPANALPARSEAAAVQPVIGISQVRNMSKEKKDLGTLGYLGITMMVI 180
QY 181 IIAIGAGIILGYSYRGKDLKEQHDQVCEREMQRTPLPSAFTNPCEIYDEKTVVHT 240
DB 181 IIAIGAGIILGYSYRGKDLKEQHDQVCEREMQRTPLPSAFTNPCEIYDEKTVVHT 240
QY 241 SQTVPDQEGSTPLMGAGTPGA 263
DB 241 SQTVPDQEGSTPLMGAGTPGA 263

RESULT 6

US-10-057-951-2
; Sequence 2 Application US/10057951
; GENERAL INFORMATION:
; APPLICANT: Moore et al.
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
; FILE REFERENCE: PF378P1
; CURRENT APPLICATION NUMBER: US/10/057,951
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US 09/411,977
; PRIOR FILING DATE: 1999-10-04
; PRIOR APPLICATION NUMBER: US 09/084,491
; PRIOR FILING DATE: 1998-05-27
; PRIOR APPLICATION NUMBER: US 60/048,000
; PRIOR FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-951-2

Query Match 100.0%; Score 1393; DB 24; Length 263;
Best Local Similarity 100.0%; Pred. No. 6,3e-129;
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLTAWQAFIVSNMLAEAYSGGCFWMDNGHLVREDQTSAPGRLCLNMLDQSGLASAP 60
DB 1 MLTAWQAFIVSNMLAEAYSGGCFWMDNGHLVREDQTSAPGRLCLNMLDQSGLASAP 60
QY 61 VSGAGNHSYCRNPDEDPGRCVYSGEAGVPEKRPCEDLRCPTTSQALPAFTTEIOEAS 120
DB 61 VSGAGNHSYCRNPDEDPGRCVYSGEAGVPEKRPCEDLRCPTTSQALPAFTTEIOEAS 120
QY 121 EGPADDEVQFAPANALPARSEAAAVQPVIGISQVRNMSKEKKDLGTLGYLGITMMVI 180
DB 121 EGPADDEVQFAPANALPARSEAAAVQPVIGISQVRNMSKEKKDLGTLGYLGITMMVI 180
QY 181 IIAIGAGIILGYSYRGKDLKEQHDQVCEREMQRTPLPSAFTNPCEIYDEKTVVHT 240
DB 181 IIAIGAGIILGYSYRGKDLKEQHDQVCEREMQRTPLPSAFTNPCEIYDEKTVVHT 240
QY 241 SQTVPDQEGSTPLMGAGTPGA 263
DB 241 SQTVPDQEGSTPLMGAGTPGA 263

RESULT 7

US-10-102-704-2
; Sequence 2 Application US/10102704
; GENERAL INFORMATION:
; APPLICANT: Moore et al.
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protein
; FILE REFERENCE: PF378P1
; CURRENT APPLICATION NUMBER: US/10/102,704

;; CURRENT FILING DATE: 2002-03-22
;; PRIOR APPLICATION NUMBER: 09/084,491
;; PRIOR FILING DATE: 1998-05-27
;; PRIOR APPLICATION NUMBER: 60/048,000
;; PRIOR FILING DATE: 1997-05-28
;; NUMBER OF SEQ ID NOS: 16
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 2
;; LENGTH: 263
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-102-704-2

Query Match 100.0%; Score 1393; DB 25; Length 263;
Best Local Similarity 100.0%; Pred. No. 6.3e-129;
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAWVOAFLVSNMLLAAYSGGCFMNDGHLVREDQTSAPAGLRCLNWLDAOSGLASAP 60
DB 1 MLAWVOAFLVSNMLLAAYSGGCFMNDGHLVREDQTSAPAGLRCLNWLDAOSGLASAP 60
QY 61 VSGAGNHSYCNPDDEPRGPMCVSGEAGVPEKRCEDLRCPETTSQALPAFTTEIQEAS 120
DB 61 VSGAGNHSYCNPDDEPRGPMCVSGEAGVPEKRCEDLRCPETTSQALPAFTTEIQEAS 120
QY 121 ESGGADEVQVFPANALPARSEAAAVQPIGISOVRMNSKEKDLGTLGYVLGITMVI 180
DB 121 ESGGADEVQVFPANALPARSEAAAVQPIGISOVRMNSKEKDLGTLGYVLGITMVI 180
QY 181 IIAIGAGIILGYSYKRGDKLEQHDQKVCEREMORITLPLSAFTNPCEIVDEKTVVHT 240
DB 181 IIAIGAGIILGYSYKRGDKLEQHDQKVCEREMORITLPLSAFTNPCEIVDEKTVVHT 240
QY 241 SOTPVDPQEGSTPLMGAGTPGA 263
DB 241 SOTPVDPQEGSTPLMGAGTPGA 263

RESULT 8
US-60-048-000-2
;; Sequence 2, Application US/60048000
;; GENERAL INFORMATION:
;; APPLICANT: EBER, REINHARD
;; APPLICANT: MOORE, PAUL
;; APPLICANT: RUBEN, STEVE
;; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE
;; NUMBER OF SEQUENCES: 16
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
;; STREET: 9410 KEY WEST AVENUE
;; CITY: ROCKVILLE
;; STATE: MD
;; COUNTRY: USA
;; ZIP: 20850
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/60/048,000
;; FILING DATE:
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BROOKS, ANDERS A.
;; REGISTRATION NUMBER: 36,373
;; REFERENCE/DOCKET NUMBER: PF379PP
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (301) 309-8504
;; TELEFAX: (301) 309-8512
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 263 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-60-048-000-2

Query Match 100.0%; Score 1393; DB 25; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.3e-128;
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAWVOAFLVSNMLLAAYSGGCFMNDGHLVREDQTSAPAGLRCLNWLDAOSGLASAP 60
DB 1 MLAWVOAFLVSNMLLAAYSGGCFMNDGHLVREDQTSAPAGLRCLNWLDAOSGLASAP 60
QY 61 VSGAGNHSYCNPDDEPRGPMCVSGEAGVPEKRCEDLRCPETTSQALPAFTTEIQEAS 120
DB 61 VSGAGNHSYCNPDDEPRGPMCVSGEAGVPEKRCEDLRCPETTSQALPAFTTEIQEAS 120
QY 121 ESGGADEVQVFPANALPARSEAAAVQPIGISOVRMNSKEKDLGTLGYVLGITMVI 180
DB 121 ESGGADEVQVFPANALPARSEAAAVQPIGISOVRMNSKEKDLGTLGYVLGITMVI 180
QY 181 IIAIGAGIILGYSYKRGDKLEQHDQKVCEREMORITLPLSAFTNPCEIVDEKTVVHT 240
DB 181 IIAIGAGIILGYSYKRGDKLEQHDQKVCEREMORITLPLSAFTNPCEIVDEKTVVHT 240
QY 241 SOTPVDPQEGSTPLMGAGTPGA 263
DB 241 SOTPVDPQEGSTPLMGAGTPGA 263

RESULT 9
US-09-791-537-39562
;; Sequence 39562, Application US/09791537
;; GENERAL INFORMATION:
;; APPLICANT: Bionomix, Inc.
;; APPLICANT: Debe, Derek
;; APPLICANT: Danzer, Joseph
;; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB.
;; FILE REFERENCE: 261/210
;; CURRENT APPLICATION NUMBER: US/09/791,537
;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 39562
;; LENGTH: 263
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-791-537-39562

Query Match 99.8%; Score 1390; DB 21; Length 263;
Best Local Similarity 99.6%; Pred. No. 1.3e-128;
Matches 262; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLAWVOAFLVSNMLLAAYSGGCFMNDGHLVREDQTSAPAGLRCLNWLDAOSGLASAP 60
DB 1 MLAWVOAFLVSNMLLAAYSGGCFMNDGHLVREDQTSAPAGLRCLNWLDAOSGLASAP 60
QY 61 VSGAGNHSYCNPDDEPRGPMCVSGEAGVPEKRCEDLRCPETTSQALPAFTTEIQEAS 120
DB 61 VSGAGNHSYCNPDDEPRGPMCVSGEAGVPEKRCEDLRCPETTSQALPAFTTEIQEAS 120
QY 121 ESGGADEVQVFPANALPARSEAAAVQPIGISOVRMNSKEKDLGTLGYVLGITMVI 180
DB 121 ESGGADEVQVFPANALPARSEAAAVQPIGISOVRMNSKEKDLGTLGYVLGITMVI 180
QY 181 IIAIGAGIILGYSYKRGDKLEQHDQKVCEREMORITLPLSAFTNPCEIVDEKTVVHT 240
DB 181 IIAIGAGIILGYSYKRGDKLEQHDQKVCEREMORITLPLSAFTNPCEIVDEKTVVHT 240
QY 241 SOTPVDPQEGSTPLMGAGTPGA 263
DB 241 SOTPVDPQEGSTPLMGAGTPGA 263

db 241 SQTPVDEBGTPLMGAGTPGA 263

```

RESULT 10
US-09-927-796-44
; Sequence 44, Application US/09927796
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Masters, Scott A.
; APPLICANT: Pan, James
; APPLICANT: Pitt, Robert M.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stone, Donna M.
; APPLICANT: Metanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
; FILE REFERENCE: F2931R1C1
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/014699
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 60/026943
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 60/059121
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/062037
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/067411
; PRIOR FILING DATE: 1997-12-03
; PRIOR APPLICATION NUMBER: 60/069862
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/095929
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/097978
; PRIOR FILING DATE: 1998-08-26
; PRIOR APPLICATION NUMBER: 60/103396
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/108867
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/112851
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/119965
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: 60/123972
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: 60/133459
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/140650
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/140653
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/144758
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145698

```

```

; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/146222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: 60/149395
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/151689
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 08/625328
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 08/710802
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 08/800699
; PRIOR FILING DATE: 1997-02-14
; PRIOR APPLICATION NUMBER: 08/826883
; PRIOR FILING DATE: 1997-03-31
; PRIOR APPLICATION NUMBER: 08/829270
; PRIOR FILING DATE: 1997-03-31
; PRIOR APPLICATION NUMBER: 08/928069
; PRIOR FILING DATE: 1997-09-11
; PRIOR APPLICATION NUMBER: 08/934494
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 09/143068
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: 09/143707
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: 09/151889
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: 09/169104
; PRIOR FILING DATE: 1998-10-09
; PRIOR APPLICATION NUMBER: 09/202089
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: 09/254311
; PRIOR FILING DATE: 1999-03-03
; PRIOR APPLICATION NUMBER: 09/304003
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: 09/380137
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/380138
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/380139
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/403297
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 09/423844
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 09/51133
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 09/511631
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 09/664610
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 09/665350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 09/690169
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 09/690189
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 09/709238
; PRIOR FILING DATE: 2000-11-18
; PRIOR APPLICATION NUMBER: 09/866034
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 09/872035
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 09/884733
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: 09/886342
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: 09/866028
; PRIOR FILING DATE: 2001-08-25
; PRIOR APPLICATION NUMBER: 09/866030
; PRIOR FILING DATE: 1997-03-31
; PRIOR APPLICATION NUMBER: 09/866030
; PRIOR FILING DATE: 1998-09-14

```

```

? PRIOR APPLICATION NUMBER: PCT/US98/19330
? PRIOR FILING DATE: 1998-09-16
? PRIOR APPLICATION NUMBER: PCT/US98/21407
? PRIOR FILING DATE: 1998-10-09
? PRIOR APPLICATION NUMBER: PCT/US98/25108
? PRIOR FILING DATE: 1998-12-01
? PRIOR APPLICATION NUMBER: PCT/US99/05026
? PRIOR FILING DATE: 1999-03-08
? PRIOR APPLICATION NUMBER: PCT/US99/12252
? PRIOR FILING DATE: 1999-06-02
? PRIOR APPLICATION NUMBER: PCT/US99/20111
? PRIOR FILING DATE: 1999-09-01
? PRIOR APPLICATION NUMBER: PCT/US99/21090
? PRIOR FILING DATE: 1999-09-15
? PRIOR APPLICATION NUMBER: PCT/US99/28313
? PRIOR FILING DATE: 1999-11-30
? PRIOR APPLICATION NUMBER: PCT/US99/28301
? PRIOR FILING DATE: 1999-12-01
? PRIOR APPLICATION NUMBER: PCT/US99/28634
? PRIOR FILING DATE: 1999-12-01
? PRIOR APPLICATION NUMBER: PCT/US99/28565
? PRIOR FILING DATE: 1999-12-02
? PRIOR APPLICATION NUMBER: PCT/US00/00219
? PRIOR FILING DATE: 2000-01-05
? PRIOR APPLICATION NUMBER: PCT/US00/03565
? PRIOR FILING DATE: 2000-02-11
? PRIOR APPLICATION NUMBER: PCT/US00/04341
? PRIOR FILING DATE: 2000-02-18
? PRIOR APPLICATION NUMBER: PCT/US00/04342
? PRIOR FILING DATE: 2000-02-18
? PRIOR APPLICATION NUMBER: PCT/US00/04414
? PRIOR FILING DATE: 2000-02-22
? PRIOR APPLICATION NUMBER: PCT/US00/05841
? PRIOR FILING DATE: 2000-03-02
? PRIOR APPLICATION NUMBER: PCT/US00/06884
? PRIOR FILING DATE: 2000-03-15
? PRIOR APPLICATION NUMBER: PCT/US00/08439
? PRIOR FILING DATE: 2000-03-30
? PRIOR APPLICATION NUMBER: PCT/US00/13705
? PRIOR FILING DATE: 2000-05-17
? PRIOR APPLICATION NUMBER: PCT/US00/14941
? PRIOR FILING DATE: 2000-05-30
? PRIOR APPLICATION NUMBER: PCT/US00/15264
? PRIOR FILING DATE: 2000-06-02
? PRIOR APPLICATION NUMBER: PCT/US00/20710
? PRIOR FILING DATE: 2000-07-28
? PRIOR APPLICATION NUMBER: PCT/US00/32678
? PRIOR FILING DATE: 2000-12-01
? PRIOR APPLICATION NUMBER: PCT/US01/17800
? PRIOR FILING DATE: 2001-06-01
? PRIOR APPLICATION NUMBER: PCT/US01/19692
? PRIOR FILING DATE: 2001-06-20
? PRIOR APPLICATION NUMBER: PCT/US01/21066
? PRIOR FILING DATE: 2001-06-29
? PRIOR APPLICATION NUMBER: PCT/US01/21735
? PRIOR FILING DATE: 2001-07-09
? NUMBER OF SEQ ID NOS: 258
? SEQ ID NO 44
? LENGTH: 263
?
Query Match          99.8%: Score 1390; DB 23; Length 263;
Best Local Similarity 99.6%: Pred. No. 1.3e-128;
Matches 262; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```

?
?
?
Qy 1 MILAWOAFVSNMLAEAYSGGCFMNDGHLREDOGTSAPAGLRCCLNWLMDAOSGLASAP 60
Db 1 MILAWOAFVSNMLAEAYSGGCFMNDGHLREDOGTSAPAGLRCCLNWLMDAOSGLASAP 60
Qy 61 VSGAGNHSYCRNPDEDPRGPMCYVSGEAGVEKRPCEBULRCPETTSQALPAFTTEIOEAS 120
Db 61 VSGAGNHSYCRNPDEDPRGPMCYVSGEAGVEKRPCEBULRCPETTSQALPAFTTEIOEAS 120
Qy 121 BCPGAEVQVFAFANALPARSEAAVQVIGISQVRNMSKEKDLGTIGYVIGITMVMVI 180
```

```

|||||
Db 121 BCPGAEVQVFAFANALPARSEAAVQVIGISQVRNMSKEKDLGTIGYVIGITMVMVI 180
Qy 181 ITAIGAGIILGYSYKRGKDLKSOHOKVCEREMORITLPLSAFTPTCEIYDEKTVVHT 240
Db 181 ITAIGAGIILGYSYKRGKDLKSOHOKVCEREMORITLPLSAFTPTCEIYDEKTVVHT 240
Qy 241 SOTPVDPQEGSTPLMGAGTPGA 263
Db 241 SOTPVDPQEGSTPLMGAGTPGA 263
```

```

RESULT 11
US-10-210-951-44
? Sequence 44, Application US/10210951
? GENERAL INFORMATION:
? APPLICANT: Ashkenazi, Avi J.
? APPLICANT: Goddard, Audrey
? APPLICANT: Godowski, Paul J.
? APPLICANT: Gurney, Austin L.
? APPLICANT: Hillan, Kenneth J.
? APPLICANT: Marselers, Scot A.
? APPLICANT: Pan, James
? APPLICANT: Pilti, Robert M.
? APPLICANT: Roy, Margaret Ann
? APPLICANT: Smith, Victoria
? APPLICANT: Stone, Donna M.
? APPLICANT: Watanabe, Colin K.
? APPLICANT: Wood, William I.
? TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
? FILE REFERENCE: P2931R1C1
? CURRENT APPLICATION NUMBER: US/10/210,951
? CURRENT FILING DATE: 2002-08-02
? PRIOR APPLICATION NUMBER: 60/014699
? PRIOR FILING DATE: 1996-04-01
? PRIOR APPLICATION NUMBER: 60/026943
? PRIOR FILING DATE: 1996-09-23
? PRIOR APPLICATION NUMBER: 60/059121
? PRIOR FILING DATE: 1997-07-17
? PRIOR APPLICATION NUMBER: 60/059352
? PRIOR FILING DATE: 1997-09-19
? PRIOR APPLICATION NUMBER: 60/062037
? PRIOR FILING DATE: 1997-10-10
? PRIOR APPLICATION NUMBER: 60/063755
? PRIOR FILING DATE: 1997-10-17
? PRIOR APPLICATION NUMBER: 60/063045
? PRIOR FILING DATE: 1997-10-24
? PRIOR APPLICATION NUMBER: 60/063046
? PRIOR FILING DATE: 1997-10-24
? PRIOR APPLICATION NUMBER: 60/065111
? PRIOR FILING DATE: 1997-11-24
? PRIOR APPLICATION NUMBER: 60/066772
? Remaining Prior Application data removed - See File Wrapper or PALM.
? NUMBER OF SEQ ID NOS: 258
? SEQ ID NO 44
? LENGTH: 263
? TYPE: PRT
? ORGANISM: Homo sapiens
US-10-210-951-44
```

```

Query Match          99.8%: Score 1390; DB 26; Length 263;
Best Local Similarity 99.6%: Pred. No. 1.3e-128;
Matches 262; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
?
?
?
Qy 1 MILAWOAFVSNMLAEAYSGGCFMNDGHLREDOGTSAPAGLRCCLNWLMDAOSGLASAP 60
Db 1 MILAWOAFVSNMLAEAYSGGCFMNDGHLREDOGTSAPAGLRCCLNWLMDAOSGLASAP 60
Qy 61 VSGAGNHSYCRNPDEDPRGPMCYVSGEAGVEKRPCEBULRCPETTSQALPAFTTEIOEAS 120
Db 61 VSGAGNHSYCRNPDEDPRGPMCYVSGEAGVEKRPCEBULRCPETTSQALPAFTTEIOEAS 120
```

```
QY 121 EGGADDEVQVFAPANALPARSEAAAVQVIGISQVRMNSKEKKDLGTLGYVLGITMMVI 180
| | | | |
DB 121 EGGADDEVQVFAPANALPARSEAAAVQVIGISQVRMNSKEKKDLGTLGYVLGITMMVI 180
QY 181 IIAIGAGIILIGSYRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIYDEKTVVHT 240
| | | | |
DB 181 IIAIGAGIILIGSYRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIYDEKTVVHT 240
QY 241 SOTPVDPQEGSTPLMGQAGTPGA 263
| | | | |
DB 241 SOTPVDPQEGSTPLMGQAGTPGA 263

RESULT 12
US-10-211-858-44
; Sequence 44, Application US/10211858
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Guiney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Masters, Scot A.
; APPLICANT: Pan, James
; APPLICANT: Pitti, Robert M.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stone, Donna M.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
; FILE REFERENCE: P2931R1C1
; CURRENT APPLICATION NUMBER: US/10/211,858
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/014699
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 60/026943
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 60/059121
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/062037
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 258
; SEQ ID NO 44
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-858-44

Query Match 99.8%; Score 1390; DB 26; Length 263;
Best Local Similarity 99.6%; Pred. No. 1.3e-128;
Matches 262; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 121 EGGADDEVQVFAPANALPARSEAAAVQVIGISQVRMNSKEKKDLGTLGYVLGITMMVI 180
| | | | |
DB 121 EGGADDEVQVFAPANALPARSEAAAVQVIGISQVRMNSKEKKDLGTLGYVLGITMMVI 180
QY 181 IIAIGAGIILIGSYRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIYDEKTVVHT 240
| | | | |
DB 181 IIAIGAGIILIGSYRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIYDEKTVVHT 240
QY 241 SOTPVDPQEGSTPLMGQAGTPGA 263
| | | | |
DB 241 SOTPVDPQEGSTPLMGQAGTPGA 263

RESULT 13
US-10-211-884-44
; Sequence 44, Application US/10211884
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Guiney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Masters, Scot A.
; APPLICANT: Pan, James
; APPLICANT: Pitti, Robert M.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stone, Donna M.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
; FILE REFERENCE: P2931R1C1
; CURRENT APPLICATION NUMBER: US/10/211,884
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/014699
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 60/026943
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 60/059121
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/062037
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 258
; SEQ ID NO 44
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-884-44

Query Match 99.8%; Score 1390; DB 26; Length 263;
Best Local Similarity 99.6%; Pred. No. 1.3e-128;
Matches 262; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```

Db      61 VSGAGNHSYCRNDEDPKPCVSGEAGVPEKPCEDLRCPETTSQALPAFTTEIOEAS 120
Qy      121 EGPDADEVQVAFAPANALPARSEAAAVQPVIGISQVRNNSKEKKDLGTLGYVLGITMVI 180
Db      121 EGPDADEVQVAFAPANALPARSEAAAVQPVIGISQVRNNSKEKKDLGTLGYVLGITMVI 180
Qy      181 IIAIGAGIILGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPTCEIVDEKTVVHT 240
Db      181 IIAIGAGIILGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPTCEIVDEKTVVHT 240
Qy      241 SQTVPDPEGSTPLMGAGTPGA 263
Db      241 SQTVPDPEGSTPLMGAGTPGA 263

```

RESULT 14

```

US-60-230-435-1580
; Sequence 1580, Application US/60230435
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL000768
; CURRENT APPLICATION NUMBER: US/60/230,435
; CURRENT FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 2991
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1580
; LENGTH: 263
; TYPE: PRP
; ORGANISM: HUMAN
US-60-230-435-1580

```

```

Query Match          99.8%; Score 1390; DB 27; Length 263;
Best Local Similarity 99.6%; Pred. No. 1.3e-128;
Matches 262; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MLAMVOAFIVSNMLAEAYSGGCFMNDNGHLYREDQTSAPAGLRCLNWLDAQSGLASAP 60
Db      1 MLAMVOAFIVSNMLAEAYSGGCFMNDNGHLYREDQTSAPAGLRCLNWLDAQSGLASAP 60
Qy      61 VSGAGNHSYCRNDEDPKPCVSGEAGVPEKPCEDLRCPETTSQALPAFTTEIOEAS 120
Db      61 VSGAGNHSYCRNDEDPKPCVSGEAGVPEKPCEDLRCPETTSQALPAFTTEIOEAS 120
Qy      121 EGPDADEVQVAFAPANALPARSEAAAVQPVIGISQVRNNSKEKKDLGTLGYVLGITMVI 180
Db      121 EGPDADEVQVAFAPANALPARSEAAAVQPVIGISQVRNNSKEKKDLGTLGYVLGITMVI 180
Qy      181 IIAIGAGIILGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPTCEIVDEKTVVHT 240
Db      181 IIAIGAGIILGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPTCEIVDEKTVVHT 240
Qy      241 SQTVPDPEGSTPLMGAGTPGA 263
Db      241 SQTVPDPEGSTPLMGAGTPGA 263

```

```

RESULT 15
US-60-389-987-244
; Sequence 244, Application US/60389987
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Bojn D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465P2

```

```

; CURRENT APPLICATION NUMBER: US/60/389,987
; CURRENT FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 3025
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 244
; LENGTH: 263
; TYPE: PRP
; ORGANISM: Homo sapiens
US-60-389-987-244

```

```

Query Match          99.8%; Score 1390; DB 27; Length 263;
Best Local Similarity 99.6%; Pred. No. 1.3e-128;
Matches 262; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MLAMVOAFIVSNMLAEAYSGGCFMNDNGHLYREDQTSAPAGLRCLNWLDAQSGLASAP 60
Db      1 MLAMVOAFIVSNMLAEAYSGGCFMNDNGHLYREDQTSAPAGLRCLNWLDAQSGLASAP 60
Qy      61 VSGAGNHSYCRNDEDPKPCVSGEAGVPEKPCEDLRCPETTSQALPAFTTEIOEAS 120
Db      61 VSGAGNHSYCRNDEDPKPCVSGEAGVPEKPCEDLRCPETTSQALPAFTTEIOEAS 120
Qy      121 EGPDADEVQVAFAPANALPARSEAAAVQPVIGISQVRNNSKEKKDLGTLGYVLGITMVI 180
Db      121 EGPDADEVQVAFAPANALPARSEAAAVQPVIGISQVRNNSKEKKDLGTLGYVLGITMVI 180
Qy      181 IIAIGAGIILGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPTCEIVDEKTVVHT 240
Db      181 IIAIGAGIILGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPTCEIVDEKTVVHT 240
Qy      241 SQTVPDPEGSTPLMGAGTPGA 263
Db      241 SQTVPDPEGSTPLMGAGTPGA 263

```

```

Search completed: April 7, 2003, 06:57:44
Job time : 195.301 secs

```

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 06:48:47 ; Search time 33.2285 Seconds

(without alignments)
1196.989 Million cell updates/sec

Title: US-10-057-951-2

1393

Sequence: 1 MLAWVQAFVSNMLAEAY.....PVDQEGSTPLMGAGTPGA 263

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 679521 seqs, 151232488 residues

Total number of hits satisfying chosen parameters: 679521

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents AA New:*
1: /cgn2_6/pdata1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/pdata1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/pdata1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/pdata1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/pdata1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/pdata1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/pdata1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1384	89.4	263	6 US-10-218-140-6002	Sequence 6002, Ap
2	1120.5	80.4	264	6 US-10-144-779-468	Sequence 488, App
3	154.5	11.1	293	7 US-60-452-680-14406	Sequence 14406, A
4	154.5	11.1	293	7 US-60-453-135-8958	Sequence 8958, Ap
5	154.5	11.1	293	7 US-60-452-680-14405	Sequence 14405, A
6	154.5	11.1	516	7 US-60-453-135-8957	Sequence 8957, Ap
7	154.5	11.1	516	7 US-60-453-135-8957	Sequence 8957, Ap
8	154.5	11.1	527	5 US-09-612-314A-51	Sequence 51, App1
9	154.5	11.1	527	5 US-09-612-314A-51	Sequence 26, App1
10	154.5	11.1	562	1 PCT-US02-3253-26	Sequence 4, App1
11	154.5	11.1	562	1 PCT-US02-3253-26	Sequence 26, App1
12	154.5	11.1	562	6 US-10-287-994-26	Sequence 26, App1
13	154.5	11.1	562	6 US-60-452-680-14408	Sequence 14408, A
14	154.5	11.1	562	7 US-60-453-135-8960	Sequence 8960, Ap
15	154.5	11.1	562	7 US-60-453-135-8960	Sequence 8960, Ap
16	154.5	11.1	587	5 US-09-949-016-11501	Sequence 11501, A
17	154.5	11.1	655	1 PCT-US02-19017-28	Sequence 28, App1
18	146.5	10.5	656	7 US-60-453-135-11516	Sequence 11516, A
19	146.5	10.5	656	7 US-60-453-135-11516	Sequence 11516, A
20	146	10.5	322	1 PCT-US02-27855-21	Sequence 21, App1
21	146	10.5	322	1 PCT-US02-27855-21	Sequence 21, App1
22	143.5	10.3	482	7 US-60-452-680-14407	Sequence 14407, A
23	143.5	10.3	482	7 US-60-453-135-8959	Sequence 8959, Ap
24	143.5	10.3	482	7 US-60-453-135-8959	Sequence 8959, Ap
25	143.5	10.3	482	7 US-60-453-135-8959	Sequence 8959, Ap
26	143	10.3	713	5 US-09-949-016-9983	Sequence 9983, Ap

27	143	10.3	713	7 US-60-452-680-22399	Sequence 22399, A
28	143	10.3	713	7 US-60-453-135-13847	Sequence 13847, A
29	143	10.3	713	7 US-60-453-050-13847	Sequence 13847, A
30	142	10.2	339	6 US-10-401-108-3	Sequence 3, App1
31	142	10.2	339	6 US-10-292-418-11	Sequence 11, App1
32	142	10.2	391	1 PCT-US02-37879-7	Sequence 7, App1
33	142	10.2	391	1 PCT-US02-37879-7	Sequence 7, App1
34	142	10.2	394	1 PCT-US02-37879-8	Sequence 8, App1
35	142	10.2	394	1 PCT-US02-37879-8	Sequence 8, App1
36	142	10.2	453	5 US-09-291-200A-1	Sequence 1, App1
37	142	10.2	458	5 US-09-946-893B-4	Sequence 4, App1
38	142	10.2	569	5 US-09-946-893B-5	Sequence 5, App1
39	142	10.2	571	5 US-09-946-893B-8	Sequence 8, App1
40	142	10.2	576	5 US-09-946-893B-6	Sequence 6, App1
41	142	10.2	791	1 PCT-US02-37879-1	Sequence 1, App1
42	142	10.2	791	1 PCT-US02-37879-1	Sequence 1, App1
43	142	10.2	810	5 US-09-291-200A-5	Sequence 5, App1
44	142	10.2	810	5 US-09-946-893B-2	Sequence 2, App1
45	141.5	10.2	688	1 PCT-US02-27855-18	Sequence 18, App1

ALIGNMENTS

```
RESULT 1
US-10-218-140-6002
; Sequence 6002, Application US/10218140
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES
; FILE REFERENCE: 15966-543 CON
; CURRENT APPLICATION NUMBER: US/10/218,140
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/127,728
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/127,636
; PRIOR FILING DATE: 1999-04-02
; PRIOR APPLICATION NUMBER: 60/127,607
; NUMBER OF SEQ ID NOS: 6122
; SOFTWARE: Curator Version 1.0
; SEQ ID NO 6002
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-218-140-6002
Query Match: 39.4%; Score 1384; DB 6; Length 263;
Best Local Similarity: 99.2%; Pred. No. 5.5e-11;
Matches 261; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MLAWVQAFVSNMLAEAYSGGCFMDNGHLRYEDQTSFAPGRRCIMYDAGSGASAP 60
Db 1 MLAWVQAFVSNMLAEAYSGGCFMDNGHLRYEDQTSFAPGRRCIMYDAGSGASAP 60
QY 61 VSGAGNHSYCRNPDDPGRPCVYSGAGVPEKRPCEDRCPETTSQALPAFTTEIOEAS 120
Db 61 VSGAGNHSYCRNPDDPGRPCVYSGAGVPEKRPCEDRCPETTSQALPAFTTEIOEAS 120
QY 121 EGPADDEVQFAPANALPARSEAAVOPVIGISQRYRMSKSKDGLTGVYGIIMVY 180
Db 121 EGPADDEVQFAPANALPARSEAAVOPVIGISQRYRMSKSKDGLTGVYGIIMVY 180
QY 181 IIAIGAGIILGYSYRGKDLKEQHDQVCEREMORTTLPISAFNPTCEIVDEKTVVHT 240
Db 181 IIAIGAGIILGYSYRGKDLKEQHDQVCEREMORTTLPISAFNPTCEIVDEKTVVHT 240
QY 241 SOTPVDPQSGSTPLMGAGTPGA 263
Db 241 SOTPVDPQSGSTPLMGAGTPGA 263
```



```
; APPLICANT: CARGILL, Michele
; APPLICANT: GRUBER, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01450
; CURRENT APPLICATION NUMBER: US/60/452,680
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14405
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-452-680-14405

Query Match
Best Local Similarity 39.6%; DB 7; Length 516;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 25 CFMDNGHLVREDQTSPPAPGLRCLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 75
DB 81 CYEDQGISYRGWTSTAESGAECTNW--NSSALAQKPYSGRRPDALRLGLGNHNYCNPDR 138
QY 76 DPRGPMCVYSGEAGVPEKRPCEDLRCPEPTS 106
DB 139 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 167

RESULT 7
US-60-453-135-8957
; Sequence 8957, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01456
; CURRENT APPLICATION NUMBER: US/60/453,135
; NUMBER OF SEQ ID NOS: 2003-03-10
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8957
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-8957

Query Match
Best Local Similarity 39.6%; DB 7; Length 516;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 25 CFMDNGHLVREDQTSPPAPGLRCLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 75
DB 81 CYEDQGISYRGWTSTAESGAECTNW--NSSALAQKPYSGRRPDALRLGLGNHNYCNPDR 138
QY 76 DPRGPMCVYSGEAGVPEKRPCEDLRCPEPTS 106
DB 139 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 167

RESULT 8
US-60-453-050-8957
; Sequence 8957, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 8957
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-8957

Query Match
Best Local Similarity 39.6%; DB 7; Length 516;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 25 CFMDNGHLVREDQTSPPAPGLRCLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 75
DB 81 CYEDQGISYRGWTSTAESGAECTNW--NSSALAQKPYSGRRPDALRLGLGNHNYCNPDR 138
QY 76 DPRGPMCVYSGEAGVPEKRPCEDLRCPEPTS 106
DB 139 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 167

RESULT 9
US-09-612-314A-51
; Sequence 51, Application US/09612314A
; GENERAL INFORMATION:
; APPLICANT: SMITH, RICHARD ANTHONY GODWIN
; APPLICANT: DODD, IAN
; APPLICANT: MOSSAKOWSKA, DANUTA EWA IRENA
; TITLE OF INVENTION: CONJUGATES OF SOLUBLE PEPTIDIC COMPOUNDS WITH
; TITLE OF INVENTION: MEMBRANE-BINDING AGENTS
; FILE REFERENCE: 37945-0004
; CURRENT APPLICATION NUMBER: US/09/612,314A
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 09/214,913
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: PCT/EP97/03715
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: GB 96 148 71.3
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: tissue plasminogen activator
US-09-612-314A-51

Query Match
Best Local Similarity 39.6%; DB 5; Length 527;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 25 CFMDNGHLVREDQTSPPAPGLRCLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 75
DB 92 CYEDQGISYRGWTSTAESGAECTNW--NSSALAQKPYSGRRPDALRLGLGNHNYCNPDR 149
QY 76 DPRGPMCVYSGEAGVPEKRPCEDLRCPEPTS 106
DB 150 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 178

RESULT 10
PCT-US02-32263-26
; Sequence 26, Application PCT/US0232263
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: Deftrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Bower, Caryn
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; TITLE OF INVENTION: REMODELING AND GLYCOCONJUGATION OF PEPTIDES
; FILE REFERENCE: 040853-01-5050WO
```

```

; CURRENT APPLICATION NUMBER: PCT/US02/32263
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/334,233
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/334,301
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-32263-26

```

```

Query Match      11.1%; Score 154.5; DB 1; Length 562;
Best Local Similarity 39.6%; Pred. No. 6,5e-05;
Matches 36; Conservative 6; Mismatches 13; Gaps 4;

```

```

Qy 25 CFWDNGHLVREDQTSAPGLRCLNWLDAQSLASAPVS-----GAGNHSYCRNPDE 75
Db 127 CYEDQGISYRGWTSTAESGAECTNW--NSSALAQKPYSGRRPDAIRLGLGNHNYCNPDR 184

```

```

Qy 76 DPRGPMCVSAGEAVPEKRPCEDLRCPEPTTS 106
Db 185 DSK-FMCIYF-KAGKYSSEFCSTPACSEGNS 213

```

```

; RESULT 11
; US-09-703-695A-4
; Sequence 4, Application US/09703695A
; GENERAL INFORMATION:
; APPLICANT: Xu, Yuan
; TITLE OF INVENTION: REVERSE-PHASE HPLC ASSAY FOR PLASMINOGEN ACTIVATORS
; FILE REFERENCE: P176881
; CURRENT APPLICATION NUMBER: US/09/703,695A
; CURRENT FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: US 60/163,607
; PRIOR FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 4
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-703-695A-4

```

```

Query Match      11.1%; Score 154.5; DB 5; Length 562;
Best Local Similarity 39.6%; Pred. No. 6,5e-05;
Matches 36; Conservative 6; Mismatches 13; Gaps 4;

```

```

Qy 25 CFWDNGHLVREDQTSAPGLRCLNWLDAQSLASAPVS-----GAGNHSYCRNPDE 75
Db 127 CYEDQGISYRGWTSTAESGAECTNW--NSSALAQKPYSGRRPDAIRLGLGNHNYCNPDR 184

```

```

Qy 76 DPRGPMCVSAGEAVPEKRPCEDLRCPEPTTS 106
Db 185 DSK-FMCIYF-KAGKYSSEFCSTPACSEGNS 213

```

RESULT 12

```

US-10-287-994-26
; Sequence 26, Application US/10287994
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: Defrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Baker, Caryn
; APPLICANT: Chen, Xi
; TITLE OF INVENTION: REMODELING AND GLYCOCONJUGATION OF PEPTIDES
; FILE REFERENCE: 040833-01-5052-00
; CURRENT APPLICATION NUMBER: US/10/287,994
; CURRENT FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-287-994-26

```

```

Query Match      11.1%; Score 154.5; DB 6; Length 562;
Best Local Similarity 39.6%; Pred. No. 6,5e-05;
Matches 36; Conservative 6; Mismatches 13; Gaps 4;

```

```

Qy 25 CFWDNGHLVREDQTSAPGLRCLNWLDAQSLASAPVS-----GAGNHSYCRNPDE 75
Db 127 CYEDQGISYRGWTSTAESGAECTNW--NSSALAQKPYSGRRPDAIRLGLGNHNYCNPDR 184

```

```

Qy 76 DPRGPMCVSAGEAVPEKRPCEDLRCPEPTTS 106
Db 185 DSK-FMCIYF-KAGKYSSEFCSTPACSEGNS 213

```

```

; RESULT 13
; US-60-452-680-14408
; Sequence 14408, Application US/60452680
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: GROPE, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CI001450
; CURRENT APPLICATION NUMBER: US/60/452,680
; CURRENT FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14408
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-452-680-14408

```

```

Query Match      11.1%; Score 154.5; DB 7; Length 562;
Best Local Similarity 39.6%; Pred. No. 6,5e-05;
Matches 36; Conservative 6; Mismatches 13; Gaps 4;

```

```

Qy 25 CFWDNGHLVREDQTSAPGLRCLNWLDAQSLASAPVS-----GAGNHSYCRNPDE 75
Db 127 CYEDQGISYRGWTSTAESGAECTNW--NSSALAQKPYSGRRPDAIRLGLGNHNYCNPDR 184

```

```

Qy 76 DPRGPMCVSAGEAVPEKRPCEDLRCPEPTTS 106
Db 185 DSK-FMCIYF-KAGKYSSEFCSTPACSEGNS 213

```

US-60-452-680-14408

```

Query Match      11.1%; Score 154.5; DB 7; Length 562;
Best Local Similarity 39.6%; Pred. No. 6,5e-05;
Matches 36; Conservative 6; Mismatches 13; Gaps 4;

```

Qy 25 CFWDNGHLVREDQTSAPGLRCLNWLDAQSLASAPVS-----GAGNHSYCRNPDE 75

Job time : 36.2285 secs

```
Db 127 CYEDOGISYRGWTWSTASGAECTNM--NSSALAQKPYSGRRDARIQLGHNHNCRNPD 184
Qy 76 DPRGPMCYVSGAGVPEKRCPCEDLRCPETTS 106
Db 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 213
```

RESULT 14

```
US-60-453-135-8960
; Sequence 8960, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 8960
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-8960
```

```
Query Match 11.1%; Score 154.5; DB 7; Length 562;
Best Local Similarity 39.6%; Pred. No. 6.5e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
```

```
Qy 25 CFWDNGHLYREDOTSPAPGRCLNWLDAQSGLASAPV-----GAGNHSYCRNPDE 75
Db 127 CYEDOGISYRGWTWSTASGAECTNM--NSSALAQKPYSGRRDARIQLGHNHNCRNPD 184
Qy 76 DPRGPMCYVSGAGVPEKRCPCEDLRCPETTS 106
Db 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 213
```

RESULT 15

```
US-60-453-050-8960
; Sequence 8960, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 8960
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-8960
```

```
Query Match 11.1%; Score 154.5; DB 7; Length 562;
Best Local Similarity 39.6%; Pred. No. 6.5e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
```

```
Qy 25 CFWDNGHLYREDOTSPAPGRCLNWLDAQSGLASAPV-----GAGNHSYCRNPDE 75
Db 127 CYEDOGISYRGWTWSTASGAECTNM--NSSALAQKPYSGRRDARIQLGHNHNCRNPD 184
Qy 76 DPRGPMCYVSGAGVPEKRCPCEDLRCPETTS 106
Db 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 213
```

Search completed: April 7, 2003, 07:03:15

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 06:59:57 ; Search time 15.2003 Seconds
(without alignments)
509.085 Million cell updates/sec

Title: US-10-057-951-2
Perfect score: 1393
Sequence: 1 MLAWQAFIVSNMLAEVY.....PVDFQSGSTPLMGQAGTPGA 263

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgnt_6/prodata/1/aa/5A.COMB.pep:*
2: /cgnt_6/prodata/1/aa/5B.COMB.pep:*
3: /cgnt_6/prodata/1/aa/6A.COMB.pep:*
4: /cgnt_6/prodata/1/aa/6B.COMB.pep:*
5: /cgnt_6/prodata/1/aa/PTUS.COMB.pep:*
6: /cgnt_6/prodata/1/aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1393	100.0	263	4	US-09-411-977-2 Sequence 2, Appl1
2	153.5	11.7	422	2	US-08-811-949-63 Sequence 63, Appl1
3	154.5	11.1	437	2	US-08-811-949-49 Sequence 49, Appl1
4	154.5	11.1	437	2	US-08-811-949-51 Sequence 51, Appl1
5	154.5	11.1	437	2	US-08-811-949-55 Sequence 55, Appl1
6	154.5	11.1	437	2	US-08-811-949-57 Sequence 57, Appl1
7	154.5	11.1	327	1	US-07-609-5108-16 Sequence 16, Appl1
8	154.5	11.1	327	2	US-08-811-949-39 Sequence 39, Appl1
9	154.5	11.1	327	2	PCT-US91-01025A-2 Sequence 2, Appl1
10	154.5	11.1	327	6	5185259-8 Patent No. 5185259
11	154.5	11.1	346	6	5520913-1 Patent No. 5520913
12	154.5	11.1	346	6	5200340-6 Patent No. 5200340
13	154.5	11.1	362	2	US-08-811-949-43 Sequence 43, Appl1
14	154.5	11.1	362	2	US-08-560-098A-50 Sequence 50, Appl1
15	154.5	11.1	362	2	US-08-885-795A-38 Sequence 38, Appl1
16	154.5	11.1	362	6	5185259-3 Patent No. 5185259
17	154.5	11.1	362	6	5200340-2 Patent No. 5200340
18	154.5	11.1	362	6	5344773-2 Patent No. 5344773
19	148.5	10.7	83	2	US-08-811-949-2 Sequence 2, Appl1
20	146.5	10.5	655	1	US-08-148-910-12 Sequence 12, Appl1
21	146.5	10.5	655	1	US-08-448-937A-12 Sequence 12, Appl1
22	142	10.2	160	3	US-08-612-788-35 Sequence 35, Appl1
23	142	10.2	160	3	US-09-066-028-35 Sequence 35, Appl1
24	142	10.2	250	3	US-08-612-788-30 Sequence 30, Appl1
25	142	10.2	250	3	US-09-066-028-30 Sequence 30, Appl1
26	142	10.2	339	1	US-08-248-629A-3 Sequence 3, Appl1
27	142	10.2	339	1	US-08-451-932-3 Sequence 3, Appl1

28	142	10.2	339	1	US-08-452-260-3 Sequence 3, Appl1
29	142	10.2	339	2	US-08-325-785-3 Sequence 3, Appl1
30	142	10.2	339	2	US-08-425-785-3 Sequence 3, Appl1
31	142	10.2	339	2	US-08-605-598B-3 Sequence 3, Appl1
32	142	10.2	339	2	US-08-425-743-3 Sequence 3, Appl1
33	142	10.2	339	2	US-08-886-735-3 Sequence 3, Appl1
34	142	10.2	339	2	US-09-066-028-3 Sequence 3, Appl1
35	142	10.2	339	5	PCT-US95-05107-3 Sequence 3, Appl1
36	142	10.2	352	2	US-08-612-788-40 Sequence 40, Appl1
37	142	10.2	352	2	US-09-066-028-40 Sequence 40, Appl1
38	142	10.2	374	4	US-09-377-250-3 Sequence 3, Appl1
39	142	10.2	374	4	US-09-377-250-2 Sequence 2, Appl1
40	142	10.2	378	2	US-08-612-788-42 Sequence 42, Appl1
41	142	10.2	378	3	US-09-066-028-42 Sequence 42, Appl1
42	142	10.2	378	4	US-09-206-059-1 Sequence 1, Appl1
43	142	10.2	451	4	US-09-377-250-1 Sequence 1, Appl1
44	142	10.2	452	4	US-09-377-250-4 Sequence 4, Appl1
45	142	10.2	790	1	US-08-469-486-54 Sequence 54, Appl1

ALIGNMENTS

RESULT 1
US-09-411-977-2
; Sequence 2, Application US/09411977
; Patent No. 6372473
; GENERAL INFORMATION:
; APPLICANT: Moore, Paul A.
; APPLICANT: Ruben, Steven M.
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Tissue Plasminogen Activator-like Protease
; FILE REFERENCE: PF378P1
; CURRENT APPLICATION NUMBER: US/09/411,977
; EARLIER FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: 09/084,491
; EARLIER FILING DATE: 1998-05-27
; EARLIER APPLICATION NUMBER: 60/048,000
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-411-977-2

Query Match	100.0%	Score 1393;	DB 4;	Length 263;
Best Local Similarity	100.0%	Pred. No 1.7e+139;		
Matches	263;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;
QY	1	MLAWQAFIVSNMLAEVYSGGCFWDSQSPAPGRCLNMLDQSGLASAP	60	
DB	1	MLAWQAFIVSNMLAEVYSGGCFWDSQSPAPGRCLNMLDQSGLASAP	60	
QY	61	VSAGNNSYKRNDEDPFGWCYSGAGYFEXRCEHLCAPETTSQMLPATTHIEAS	120	
DB	61	VSAGNNSYKRNDEDPFGWCYSGAGYFEXRCEHLCAPETTSQMLPATTHIEAS	120	
QY	121	BSPGADDEVQFAPANALPAREEAAVQPIYISQVRNNSYKKDKLCTGYLGITMMVY	180	
DB	121	BSPGADDEVQFAPANALPAREEAAVQPIYISQVRNNSYKKDKLCTGYLGITMMVY	180	
QY	181	IIAIGAGIIGSYKRGDLXEQHDKVCRERMOXITLPLSAFTNPTGCIIVDEKTVVHT	240	
DB	181	IIAIGAGIIGSYKRGDLXEQHDKVCRERMOXITLPLSAFTNPTGCIIVDEKTVVHT	240	
QY	241	SQTVPDQSGSTPLMGQAGTPGA 263		
DB	241	SQTVPDQSGSTPLMGQAGTPGA 263		
RESULT 2				

```

US-08-811-949-63
Sequence 63, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIWA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ. ID NO.: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 472 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-63

Query Match 11.7% Score 163.5; DB 2; Length 472;
Best Local Similarity 40.0%; Pred No. 7.2e-09;
Matches 38; Conservative 6; Mismatches 38; Indels 13; Gaps 4

Cy 21 GGGCGFNDNHLREDQTSAPRLRLNMLDAGSLASAPVS-----GGNHSYCR 71
Db 33 GARSLELDQISIRGWTWIASGNECTN--NSRLNQRFSGRRPDRILGNNHYCR 90
Cy 72 NDEDEPRGWCYCYSGEAGVPEKPKREDLRCPETTS 106
Db 91 NEDRDSK-FWCIVF--KAGKYSERFCSTPACSEKNS 123

RESULT 3
US-08-811-949-49
Sequence 49, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIWA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.

```

```

1 STREET 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
2 CITY: ARLINGTON
3 STATE: VA
4 COUNTRY: USA
5 ZIP: 22202
6 COMPUTER READABLE FORM:
7 MEDIUM TYPE: Floppy disk
8 COMPUTER: IBM PC compatible
9 OPERATING SYSTEM: PC-DOS/MS-DOS
10 SOFTWARE: Patent Release #1.0, Version #1.30
11 CURRENT APPLICATION NUMBER: US/08/811,949
12 FILING DATE: 05-MAR-1997
13 CLASSIFICATION: 435
14 ATTORNEY/AGENT INFORMATION:
15 NAME: OBOLON, NORMAN F.
16 REGISTRATION NUMBER: 24,618
17 REFERENCE/DOCKET NUMBER: 18-966-0
18 TELECOMMUNICATION INFORMATION:
19 TELEPHONE: 703-413-1000
20 TELEFAX: 703-413-2220
21 INFORMATION FOR SEQ ID NO.: 49:
22 SEQUENCE CHARACTERISTICS:
23 LENGTH: 437 amino acids
24 TYPE: amino acid
25 TOPOLOGY: linear
26 MOLECULE TYPE: Protein
27 US-08-811-949-49
28
29 Query Match 11.1%; Score 154.5; DB 2; Length 437;
30 Best Local Similarity 39.6%; Pred. No. 5,8e+08;
31 Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps
32
33 Db 76 DPROGMYTSGAGVPEKRCDELRPETTS 106
34 2 CLEQDGLSTHGTWTAESGAECLTM--NSSALNKKYSSRRPDRIQLGNINVCANPDR 59
35 1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
36 |-----GAGNSYGRNDE 75
37
38 RESULT 4
39 US-08-811-949-51
40 Sequence 51, Application US/08811949
41 Patent No. 5840531
42 GENERAL INFORMATION:
43 APPLICANT: NINA, MINBO
44 APPLICANT: SAITO, YOSHIMASA
45 APPLICANT: SASAKI, HITOSHI
46 APPLICANT: HAYASHI, MASAO
47 APPLICANT: NOTANI, JOUJI
48 APPLICANT: KOBAYASHI, MASAAZU
49 TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
50 NUMBER OF SEQUENCES: 67
51 CORRESPONDENCE ADDRESS:
52 ADDRESSEE: OBOLON, SEIVAK, MCCLELLAND, MAIER & NEUSTADT*,
53 ADDRESS: P.C.
54 STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
55 CITY: ARLINGTON
56 STATE: VA
57 COUNTRY: USA
58 ZIP: 22202
59 COMPUTER READABLE FORM:
60 MEDIUM TYPE: Floppy disk
61 COMPUTER: IBM PC compatible
62 OPERATING SYSTEM: PC-DOS/MS-DOS
63 SOFTWARE: Patent Release #1.0, Version #1.30
64 CURRENT APPLICATION DATA:
65 APPLICATION NUMBER: US/08/811,949
66 FILING DATE: 05-MAR-1997
67 CLASSIFICATION: 435
68 ATTORNEY/AGENT INFORMATION:

```


Patent No. 5326700
GENERAL INFORMATION:
APPLICANT: Berg et al.
TITLE OF INVENTION: Method for Altering Post-Translational Processing of Tissue PI
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: IN
COUNTRY: U.S.A.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/609,510B
FILING DATE: 19901106
CLASSIFICATION: 435
INFORMATION FOR SEQ. ID NO.: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-609-510B-16

Query Match 11.1%; Score 154.5; DB 1; Length 527;
Best Local Similarity 39.6%; Pred. No. 7.7e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 25 CFWDNGHLYREDQTSAPGLRCLNWLDAQSLASAPVS-----GAGNHSYCRNPDE 75
DB 92 CYEDQGISYRGWSTBESGAECTNW--NSSALAKKPYSGRRPDALRLGLGNHNYCRNPDR 149
QY 76 DPRGPMCVSGEAGVPEKRPCEDLRCPEPTTS 106
DB 150 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 178

RESULT 8
US-08-811-949-39
Sequence 39; Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIVA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ. ID NO.: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-39

Query Match 11.1%; Score 154.5; DB 2; Length 527;
Best Local Similarity 39.6%; Pred. No. 7.7e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 25 CFWDNGHLYREDQTSAPGLRCLNWLDAQSLASAPVS-----GAGNHSYCRNPDE 75
DB 92 CYEDQGISYRGWSTBESGAECTNW--NSSALAKKPYSGRRPDALRLGLGNHNYCRNPDR 149
QY 76 DPRGPMCVSGEAGVPEKRPCEDLRCPEPTTS 106
DB 150 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 178

RESULT 9
PCT-US91-01025A-2
Sequence 2; Application PC/TUS9101025A
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
TITLE OF INVENTION: Tissue Plasminogen Activator Having Fibrin
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/01025A
FILING DATE: 19910214
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/486,657
FILING DATE: 1 March 1990
ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 454P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/266-1896
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ. ID NO.: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
PCT-US91-01025A-2

Query Match 11.1%; Score 154.5; DB 5; Length 527;
Best Local Similarity 39.6%; Pred. No. 7.7e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: April 7, 2003, 08:41:02 ; Search time 19 Seconds
(without alignments)
1325.643 Million cell updates/sec

Title: US-10-057-951-2_COPY_2_263

Perfect score: 1388

Sequence: 1 LAMWQAFIVSNMLAEAYG.....PVDQEGSTPLMGQACTPGA 262

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 73:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	154.5	11.1	562	1 UKHUT	t-plasminogen acti
2	154	11.1	559	1 A35029	t-plasminogen acti
3	147.5	10.6	559	1 A29941	t-plasminogen acti
4	146.5	10.6	655	1 A46688	hepatocyte growth
5	145.5	10.5	291	2 I38098	t-plasminogen acti
6	142	10.2	810	1 PLHU	plasmin (EC 3.4.21
7	138	9.9	169	2 A40522	plasmin (EC 3.4.21
8	138	9.9	433	1 UKBAY	u-plasminogen acti
9	137	9.9	716	1 JC5061	macrophage-stimula
10	135.5	9.8	442	1 UKPG	u-plasminogen acti
11	134	9.7	431	2 US0599	t-plasminogen acti
12	134	9.7	477	2 A34369	t-plasminogen acti
13	134	9.7	477	2 JS0598	t-plasminogen acti
14	134	9.7	477	2 A40332	macrophage-stimula
15	132	9.5	394	2 JS0600	t-plasminogen acti
16	132	9.5	431	1 UKHU	u-plasminogen acti
17	130	9.4	810	2 I46260	plasmin (EC 3.4.21
18	128.5	9.3	434	1 A35005	u-plasminogen acti
19	128.5	9.1	433	1 A35005	u-plasminogen acti
20	126.5	9.1	728	1 A60185	hepatocyte growth
21	125.5	9.0	560	1 A60185	plasma hyaluronan-
22	124.5	9.0	610	2 B30848	thrombin (EC 3.4.21
23	124	8.9	622	1 TBHU	hepatocyte growth
24	124	8.9	728	1 JH0579	hepatocyte growth
25	124	8.9	812	1 PLBO	plasmin (EC 3.4.21
26	123.5	8.9	593	2 S45281	coagulation factor
27	123.5	8.8	4548	1 S00657	apoptosis factor
28	122	8.8	603	2 S28941	coagulation factor
29	121.5	8.8	728	1 A35644	hepatocyte growth

30	120	8.6	433	1 UKMS	u-Plasminogen acti
31	120	8.6	790	1 PLPG	plasmin (EC 3.4.21
32	119.5	8.6	477	2 JS0597	t-plasminogen acti
33	119	8.6	710	1 I51283	hepatocyte growth
34	118.5	8.5	615	1 KFHU12	coagulation factor
35	118	8.5	432	1 S18932	u-Plasminogen acti
36	117	8.4	625	1 TBBO	thrombin (EC 3.4.2
37	116	8.4	1420	2 A32869	apolipoprotein(a)
38	115.5	8.3	711	1 A47136	macrophage-stimula
39	114.5	8.2	812	1 PLMS	plasmin (EC 3.4.21
40	114	8.2	618	2 A35827	thrombin (EC 3.4.2
41	113	8.1	558	2 UC5878	plasma hyaluronan-
42	112	8.1	455	2 A61545	plasmin (EC 3.4.21
43	110.5	8.0	685	1 A48289	neurotrophic recep
44	108.5	7.8	123	2 C61545	plasmin (EC 3.4.21
45	107	7.7	617	2 S10511	thrombin (EC 3.4.2

ALIGNMENTS

RESULT 1

UKHUT t-plasminogen activator (EC 3.4.21.68) precursor [validated] - human

N/Alternate names: t-PA; tissue plasminogen activator

C/Species: Homo sapiens (man)

C/Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #ext-change 08-Dec-2000

C/Accession: A94004; A23529; J0562; A93293; S02125; A91343; A93951; A91322; A54645; I601

R/Ref: Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984

A/Title: The structure of the human tissue plasminogen activator gene: correlation

A/Reference number: A94004; MUID:84296137; PMID:6089198

A/Accession: A94004

A/Molecule type: DNA

A/Residues: 1-562 <NT>

A/Cross-references: GB:I00141

A/Note: The codon given for residue 93 (ACC) is inconsistent with the authors' translati

R/Ref: Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984

A/Title: Purification and characterization of tissue plasminogen activator secreted by h

A/Reference number: J0562; MUID:91291340; PMID:1368681

A/Accession: J0562

A/Molecule type: mRNA

A/Residues: 31-562 <ITA>

A/Cross-references: DDBJ:D01096; NID:g220128; PIRN:BA00881.1; PID:g441174

A/Experimental source: embryonic lung fibroblast IMR-90 cells

A/Note: part of this sequence, including the amino end of the mature protein, was confir

R/Ref: Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984

A/Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Esche

A/Reference number: A93293; MUID:83115662; PMID:6337343

A/Accession: A93293

A/Molecule type: mRNA

A/Residues: 1-562 <PEN>

A/Cross-references: GB:I00141

A/Experimental source: melanoma cells

R/Ref: Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984

A/Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human fe

A/Reference number: S02125; MUID:88262579; PMID:3133640

A/Accession: S02125

A/Molecule type: mRNA

A/Residues: 1-562 <SAS>

A/Cross-references: EMBL:X07393; NID:g37243; PIRN:CAA0302.1; PID:g37244

A;Experimental source: fetal lung cells
R;Agglutinin, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Salco, A.; Bando, H.; Okada, K.; Ma
FBS Lett. 189, 145-149, 1985
A;Title: Expression in Escherichia coli of finger-domain lacking tissue-type plasminogen
A;Reference number: A91343; PMID:85285620; PMID:3896853
A;Accession: A91343
A;Molecule type: mRNA
A;Residues: 1-38, 'G', '86-433', 'E', '435-562' <RAG>
A;Experimental source: AtCC 138
R;Lund, T.; Ny, T.; Ranby, M.; Heden, L.O.; Palm, G.; Holmgren, E.; Josephson, S.
Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983
A;Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen acti
A;Reference number: A93951; PMID:83196566; PMID:6572897
A;Accession: A93951
A;Molecule type: mRNA
A;Residues: 251-358 <BD>
A;Experimental source: melanoma cells
R;Polh, G.; Kallstrom, M.; Bergsdorf, N.; Wallen, P.; Jorvall, H.
Biochemistry 23, 3701-3707, 1984
A;Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived am
differences.
A;Reference number: A90489; PMID:85000468; PMID:6433976
A;Contents: annotation; melanoma cells, partial sequence of residues 36-562, active and
R;Polh, G.; Kaplan, L.; Einarsson, M.; Wallen, P.; Jorvall, H.
FBS Lett. 168, 29-32, 1984
A;Title: Differences between uterine and melanoma forms of tissue plasminogen activator.
A;Reference number: A91322; PMID:84158956; PMID:6538514
A;Accession: A91322
A;Molecule type: protein
A;Residues: 33-45/311-320 <POH>
A;Experimental source: uterus
A;Note: In the uterus, cleavage of the activation peptide may also occur after 38-Gln
R;van Zonneveld, A.J.; Veerman, H.; Pannekoek, H.
J. Biol. Chem. 261, 14214-14218, 1986
A;Reference number: A37567; PMID:8703611; PMID:3021732
A;Contents: annotation; fibrin binding site
R;Verheijen, J.H.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, P.H.; Enger
EMBO J. 5, 3525-3530, 1986
A;Title: Involvement of finger domain and kringle 2 domain of tissue-type plasminogen ac
A;Reference number: A37568; PMID:87161761; PMID:3030730
A;Contents: annotation; fibrin binding site
R;Dodd, I.; Nunn, B.; Robinson, J.H.
Thromb. Haemost. 59, 523-528, 1988
A;Title: Isolation, identification and pharmacokinetic properties of human tissue-type F
A;Reference number: A60902; PMID:8904681; PMID:3142086
A;Contents: annotation; novel forms of expressed recombinant t-PA
R;Harris, T.J.R.; Patel, T.; Marston, F.A.O.; Little, S.; Emrage, J.S.; Opdenaker, G.;
Mol. Biol. Med. 3, 279-292, 1986
A;Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its exp
A;Reference number: A54645; PMID:86284200; PMID:3090401
A;Accession: A54645
A;Molecule type: mRNA
A;Residues: 1-562 <HAR>
A;Cross-references: GB:M15518; NID:9190031; PIDN:AAA60111.1; PID:9190032
A;Note: Parts of this sequence were confirmed by peptide sequencing
R;Reddy, V.B.; Garraway, A.J.; Sasak, H.; Wei, C.
DNA 6, 461-472, 1987
A;Title: Expression of human uterine tissue-type plasminogen activator in mouse cells us
A;Reference number: A60110; PMID:88054470; PMID:2824147
A;Accession: A60110
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-362 <RES>
A;Cross-references: GB:M18182; NID:9340176; PIDN:AAA36800.1; PID:9340177
R;Fisher, R.; Waller, E.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schlenning, W.D.
J. Biol. Chem. 260, 11223-11230, 1985
A;Title: Isolation and characterization of the human tissue-type plasminogen activator B
A;Reference number: I55232; PMID:85289338; PMID:3161893
A;Accession: I55232
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-36 <RE2>
A;Cross-references: GB:M1890; NID:9339837; PIDN:AAA61213.1; PID:9339839

C;Comment: Cleavage by plasmin or trypsin produces two chains held together by a single
C;Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val bond. It
C;Comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repeat
C;Gene: GDB:PLAT
A;Cross-references: GDB:119496; OMIM:173370
A;Map position: 8p12-8p12
A;Intons: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2; 297/1; 362/2; 408/1; 455/3; 5
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase
F;123/Domain: signal sequence #status predicted <SIG>
F;24-32/Domain: propeptide #status predicted <PRO>
F;33-552/Product: t-plasminogen activator #status experimental <MAT>
F;33-310/Product: t-plasminogen activator chain A #status experimental <ACH>
F;41-78/Domain: fibronectin type I repeat homology <1F1>
F;86-119/Domain: EGF homology <EGF>
F;127-208/Domain: kringle homology <KR1>
F;215-296/Domain: kringle homology <KR2>
F;311-562/Product: t-plasminogen activator chain B #status experimental <BCH>
F;311-556/Domain: trypsin homology <TRY>
F;41-71-69-78-86-97-91-108-110-119-127-208-148-190-179-203-215-296-236-278-267-291-299-
F;152/483/Binding site: carboxylate (Asn) (covalent) #status experimental
F;129/Binding site: carboxylate (Asn) (covalent) (partial) #status experimental
F;310-311/Cleavage site: Arg-Ile (plasmin, trypsin) #status experimental
F;357/406/Active site: His, Asp #status predicted
F;513/Active site: Ser #status experimental
Query Match 11.1%; Score 154.5; DB 1; Length 562;
Best Local Similarity 39.6%; Pred. No. 3.4e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
Qy 24 CFMDNGHLYREQGTSPAGRLCLNLDAGSLASAPVS-----GAGHNSCRPPE 74
Db 127 CYBQGGSTRGVSFASGAGACTNW--NSSALQKPYSGRRPDAIRLGLGNHNCRNPR 184
Qy 75 DPRGWCYVSGAGVPEKRPCEDLRCPTTS 105
Db 185 DSK-PMCYVF-RAGKYSSEFCSTPACSEGNS 213
RESULT 2
A35029
t-plasminogen activator (EC 3.4.21.68) precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A35029; A31597
R;Eng, P.; Ohlsson, M.; Ny, T.
J. Biol. Chem. 265, 2022-2027, 1990
A;Title: The structure of the rat t-PA-less rat tissue-type plasminogen activator gene. Spec
A;Reference number: A35029; PMID:9010448; PMID:2105315
A;Accession: A35029
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-559 <FEN>
A;Cross-references: GB:M11197; NID:9207429; PIDN:AAA42261.1; PID:9207431; GB:J05226
R;Ny, T.; Leondardsson, G.; Haub, A.J.W.
DNA 7, 671-677, 1988
A;Title: Cloning and characterization of a cDNA for rat tissue-type plasminogen activato
A;Reference number: A31597; PMID:89170114; PMID:3148445
A;Accession: A31597
A;Molecule type: mRNA
A;Residues: 1-379, 'K', '381-559' <NYT>
A;Cross-references: GB:M23697; NID:9530159; PIDN:AAA41812.1; PID:9530160
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;11-17/Domain: signal sequence #status predicted <SIG>
F;18-29/Domain: propeptide #status predicted <PRO>
F;30-559/Product: t-plasminogen activator #status predicted <MAT>
F;30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
F;38-75/Domain: fibronectin type I repeat homology <1F1>
F;83-116/Domain: EGF homology <EGF>
F;124-205/Domain: kringle homology <KR1>
F;213-294/Domain: kringle homology <KR2>

F:309-559/Domain: t-plasminogen activator chain B #status predicted <BCH>
F:309-553/Domain: trypsin homology <TRY>
F:38-68-66-75-83-94-88-105-107-116-124-205-145-187-176-200-213-294-234-276-265-289-297-4
F:149-461/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:308-309/Cleavage site: Arg-116 (plasmin, trypsin) #status predicted
F:355-404-510/Active site: His, Asp, Ser #status predicted

Query Match 11.1% Score 154; DB 1; Length 559;
Best Local Similarity 32.4%; Pred. No. 3.7e-05;

Matches 45; Conservative 14; Mismatches 54; Indels 26; Gaps 6;

QY 24 CFMNGHLYREDQTSPPAPGLRCLNMLDAQSGIASAPVS-----GAGNHSYCRNPDE 74
DB 124 CFEEGGITVGTWSTAEINAEICINW--NSSALSGKPYSPARPAKIKLGNNYCRNPDR 181

QY 75 DPGRCWYVSGEAGVEKPCEDLRCPE-----TTSQALPAFTTEIOESGPG 123
DB 182 DPK-PWCTYF-KGKTYTTEFCSTPACPKGPTEDCYGKGYTRGHSFTT--SKASCLPW 237

QY 124 ADEVQVAPANALPASEA 142

DB 238 NSMTLIGKTYTAWRANSOA 256

RESULT 3

t-plasminogen activator (EC 3.4.21.68) precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: A29941; S48205; S48207

R:Ricklefs, R.J.; Darlow, A.L.; Strickland, S.

J. Biol. Chem. 263, 1563-1569, 1988

A:Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mRNA

A:Reference number: A29941; MUID:86087303; PMID:2826484

A:Accession: A29941

A:Molecule type: mRNA

A:Residues: 1-53 <RIG>

A:Cross-references: GB:U03520; NID:G202109; PIND:AAA0470.1; PID:G202110

R:Liinen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.

Eur. J. Biochem. 224, 863-871, 1994

A:Title: Characterization of the murine plasma fibrinolytic system.

A:Reference number: S48202; MUID:95010076; PMID:7523120

A:Accession: S48205

A:Molecule type: protein

A:Residues: 33-37, 'X', '39-40 <LIU>

A:Accession: S48207

A:Molecule type: protein

A:Residues: 309-316 <LIU>

A:Accession: S48206

A:Molecule type: protein

A:Residues: 33-37, 'X', '39-40 <LIU>

C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom

C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-29/Domain: propeptide #status predicted <PRO>

F:30-559/Product: t-plasminogen activator #status predicted <MAT>

F:308-309/Product: t-plasminogen activator chain A #status predicted <ACH>

F:38-75/Domain: fibronectin type I repeat homology <IFI>

F:83-116/Domain: EGF homology <EGF>

F:124-205/Domain: kringle homology <KR1>

F:213-294/Domain: kringle homology <KR2>

F:309-559/Product: t-plasminogen activator chain B #status predicted <BCH>

F:38-68-66-75-83-94-88-105-107-116-124-205-145-187-176-200-213-294-234-276-265-289-297-4

F:149-461/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:308-309/Cleavage site: Arg-116 (plasmin, trypsin) #status predicted

F:355-404-510/Active site: His, Asp, Ser #status predicted

Query Match 10.6% Score 147.5; DB 1; Length 559;
Best Local Similarity 37.0%; Pred. No. 0.00013;

Matches 34; Conservative 11; Mismatches 34; Indels 13; Gaps 4;

QY 24 CFMNGHLYREDQTSPPAPGLRCLNMLDAQSGIASAPVS-----GAGNHSYCRNPDE 74

DB 124 CFEEGGITVGTWSTAEINAEICINW--NSSALSGKPYSPARPAKIKLGNNYCRNPDR 181
QY 75 DPGRCWYVSGEAGVEKPCEDLRCPE-----TTSQALPAFTTEIOESGPG 106
DB 182 DPK-PWCTYF-KGKTYTTEFCSTPACPKGPTEDCYGKGYTRGHSFTT--SKASCLPW 211

RESULT 4

hepatocyte growth factor activator (EC 3.4.21.-) precursor [validated] - human

A:Accession: A46688

C:Species: Homo sapiens (man)

C:Date: 21-Sep-1993 #sequence_revision 25-Aug-1995 #text_change 08-Dec-2000

C:Accession: A46688

R:Wiyazawa, K.; Shimomura, T.; Kitamura, A.; Kondo, J.; Morimoto, Y.; Kitamura, N.

J. Biol. Chem. 268, 10024-10028, 1993

A:Title: Molecular cloning and sequence analysis of the cDNA for a human serine protease

d coagulation factor XII.

A:Reference number: A46688; MUID:93252878; PMID:768365

A:Accession: A46688

A:Molecule type: mRNA

A:Residues: 1-655 <MTY>

A:Cross-references: DDBJ:U14012; NID:G219680; PIND:BA03113.1; PID:G219681

A:Experimental source: liver (mRNA); serum (protein)

A:Note: sequence extracted from NCBI backbone (NCBI:131227, NCBI:131228)

A:Note: parts of the sequence, including the amino ends of the heavy and light chains, cc

C:Genetics:

A:Gene: GDB:HGPAC; HGFA; HGFAF

A:Cross-references: GDB:1954514

A:Map position: 4p16-4p16

C:Function:

A:Description: activates hepatocyte growth factor by specific proteolytic cleavage

A:Superfamily: tissue repair and regeneration

C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;

C:Keywords: glycoprotein; hydrolase; kringle; liver; plasma; serine proteinase

F:1-3/Domain: signal sequence #status predicted <SIG>

F:108-148/Domain: fibronectin type II repeat homology <IF2>

F:164-197/Domain: EGF homology <EG1>

F:202-237/Domain: EGF homology <EG2>

F:245-278/Domain: EGF homology <EG3>

F:286-367/Domain: kringle homology <KRG>

F:373-407/Product: hepatocyte growth factor activator light chain #status experimental <L

F:408-655/Product: hepatocyte growth factor activator heavy chain #status experimental <H

F:408-641/Domain: trypsin homology <TRY>

F:40-48-290-468-492-546/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:164-175-169-186-188-197-202-230-228-237-245-256-250-267-269-278-286-367-307-349-338-36

F:447-497-598/Active site: His, Asp, Ser #status predicted

Query Match 10.6% Score 146.5; DB 1; Length 655;
Best Local Similarity 36.9%; Pred. No. 0.00019;

Matches 41; Conservative 10; Mismatches 41; Indels 19; Gaps 6;

QY 24 CFMNGHLYREDQTSPPAPGLRCLNMLDAQSGIASAPVSAGAGNHSYCRNPDE 75
DB 286 CFMNGHLYREDQTSPPAPGLRCLNMLDAQSGIASAPVSAGAGNHSYCRNPDE 344

QY 76 DPGRCWYVSGEAGVP-----EKPCEDLRCPEFTTSQALPAFTTEIOE-ASEG 121
DB 345 ER-PWCYVVDKSLMWEYCRLEKCSL-----TRVQLSPDLATLPSPASPG 390

RESULT 5

I38098

t-plasminogen activator precursor, inactive endothelial splice form - human

N:Alternate names: tissue plasminogen activator

C:Species: Homo sapiens (man)

C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999

C:Accession: I38098; S01678

R:Stepper, F.D.; Fong, K.

Nucleic Acids Res. 18, 1086, 1990

A:Title: Variant tissue plasminogen activator (PLAT) cDNA obtained from human endo

A:Reference number: I38098; MUID:90192128; PMID:1969145

A:Accession: I38098

A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-291 <SITE>
A/Cross-references: EMBL:X13097; NID:G35282; P1DN:CAA31489.1; PID:G35283
A/Comment: For the main splice form, see PIR:UNQUT. This form probably does not have ptc
C/Genetics:
A/Genes: GDB:PLAT
A/Cross-references: GDB:119496; OMIM:173370
A/Map position: 8p12-8p12
A/Introns: 24/3; 39/1; 85/1; 122/1; 211/1; 268/2
C/Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C/Keywords: alternative splicing; fibronolysis; glycoprotein; kringie
F/1-23/Domain: signal sequence #status predicted <SIC>
F/24-32/Domain: propeptide #status predicted <PRO>
F/33-291/Product: C-plasminogen activator, inactive endothelial splice form #status pred
F/1-78/Domain: fibronectin type I repeat homology <EFP>
F/127-208/Domain: EGF homology <EGF>
F/215-291/Domain: kringie homology <KR1>
F/41-71,69-78,86-97,91-108,110-119,127-208,148-190,179-203/Disulfide bonds: #status pred
Query Match 10.5%; Score 145.5; DB 2; Length 291;
Best Local Similarity 38.5%; Pred. No. 9e-05;
Matches 35; Conservative 6; Mismatches 37; Indels 13; Gaps 4;
QY 24 CPMNDGHLTYREDTSPAPGLRCLNMLDQSGIAPVS-----GAGNHSYCRNPDE 74
DB 127 CYEDGCIETRGWSTASGAECTNW--NSSLAQNAVSGRRPALTLGHNHYCRNPFDR 184
QY 75 DPGWCVSGEAGVPEKRCEDLACPEPTS 105
DB 165 DSK-PMCTVF-KAGKYSSEFCSTPACSENS 213

RESULT 6
PLHU
plasmin (EC 3.4.21.7) precursor [validated] - human
N/Alternate names: plasminogen precursor [misnomer]
N/Contains: angiotensin, microplasmin, plasminogen
C/Species: Homo sapiens (man)
C/Date: 24-Apr-1984 #sequence revision 02-Dec-1994 #text change 15-Sep-2000
C/Accession: A35229; I52242; A26646; I62738; I64609; S03735; A00929; A04627; A04625; A04
R/Petersen, T.E.; Martzen, M.R.; Ichinose, A.; Davie, E.W.
J. Biol. Chem. 265, 6104-6111, 1990
A/Title: Characterization of the gene for human plasminogen, a key proenzyme in the fibr
A/Reference number: A35229; MUID:50202879; PMID:2318848
A/Accession: A35229
A/Molecule type: DNA
A/Residues: 1-810 <PEPT>
A/Cross-references: GB:J05286; GB:M34276; NID:G190064; P1DN:AAA60113.1; PID:G387026
R/Experimental source: Leukocyte lung fibroblast
R/Maladrecty, N.; Bruno, L.; Pontoglio, M.; Candiani, G.; Meroni, G.; Ottolenghi, S.; Te
Biochem. Biophys. Res. Commun. 173, 1013-1018, 1990
A/Title: Definition of the transcription initiation site of human plasminogen gene in 14
A/Reference number: I52242; MUID:91097523; PMID:2268308
A/Accession: I52242
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-16 <MAL1>
A/Cross-references: GB:M62890; NID:G190092; P1DN:AAA6454.1; PID:G55313
R/Forgren, M.; Raden, B.; Israelsson, M.; Larsson, K.; Heden, L.O.
FEBS Lett. 213, 254-260, 1987
A/Title: Molecular cloning and characterization of a full-length cDNA clone for human pl
A/Reference number: A26646; MUID:87162490; PMID:3030613
A/Accession: A26646
A/Molecule type: mRNA
A/Residues: 1-471; D, 473-810 <FOR>
A/Cross-references: GB:X05199; NID:G35530; P1DN:CAA28831.1; PID:G35531
R/Experimental source: liver
R/Malinowski, D.P.; Sadler, J.E.; Davie, E.W.
Biochemistry 23, 4243-4250, 1984
A/Title: Characterization of a complementary deoxyribonucleic acid coding for human and
A/Reference number: I45961; MUID:85023311; PMID:6148961

A/Accession: I62738
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 292-471; D, 473-810 <MAL2>
A/Cross-references: GB:K02922; NID:G190112; P1DN:AAA60124.1; PID:G387031
A/Accession: I64609
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 367-419 <MAL3>
A/Cross-references: GB:K02921; NID:G190110; P1DN:AAA60123.1; PID:G190111
R/Brunnholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, B.E.; Leger, M.; Manneberg, M.
Eur. J. Biochem. 116, 465-470, 1981
A/Title: Comparison of the primary structure of the N-terminal CNBr fragments of human,
A/Reference number: S03735; MUID:8121097; PMID:7238497
A/Accession: S03735
A/Molecule type: protein
A/Residues: 20-71, 'E', '73-76 <BRU>
R/Sottrop-Jensen, L.; Petersen, T.E.; Magnusson, S.
submitted to the Atlas, July 1977
A/Reference number: A00929
A/Accession: A00929
A/Molecule type: protein
A/Residues: 20-71, 'E', '73-85, 87-106, 'D', 108-360, 'E', 362-810 <SOT>
R/Wiman, B.
Eur. J. Biochem. 76, 129-137, 1977
A/Title: Primary structure of the B-chain of human plasmin.
A/Reference number: A04627; MUID:77225245; PMID:142009
A/Accession: A04627
A/Molecule type: protein
A/Residues: 581-810 <WIL>
R/Wiman, B.; Wallen, P.
Eur. J. Biochem. 50, 489-494, 1975
A/Title: Structural relationship between "glutamic acid" and "lysine" forms of human pla
A/Reference number: A04625; MUID:75093329; PMID:122932
A/Accession: A04625
A/Molecule type: protein
A/Residues: 20-50, 'Q', '51-71, 'E', '73-85, 87-100 <W12>
R/Wiman, B.; Wallen, P.
Eur. J. Biochem. 58, 539-547, 1975
A/Title: Amino-acid sequence of the cyanogen-bromide fragment from human plasminogen that
A/Reference number: A04626; MUID:76043692; PMID:126663
A/Accession: A04626
A/Molecule type: protein
A/Residues: 483-507, 'E', 509-604 <W13>
R/Robbins, K.C.; Bernade, P.; Arzodon, L.; Summaria, L.
J. Biol. Chem. 248, 1631-1633, 1973
A/Title: The primary structure of human plasminogen. II. The histidine loop of human pl
A/Reference number: A92125; MUID:73149248; PMID:4694729
A/Contents: annotation; active site
R/Groskopf, W.R.; Summaria, L.; Robbins, K.C.
J. Biol. Chem. 244, 3590-3597, 1969
A/Title: Studies on the active center of human plasmin. Partial amino acid sequence of a
A/Reference number: A92048; MUID:69234739; PMID:4240117
A/Contents: annotation; active site
R/Trexler, M.; Valli, Z.; Pecthy, L.
J. Biol. Chem. 257, 7401-7406, 1982
A/Title: Structure of the omega-aminocarboxylic acid-binding sites of human plasminogen.
A/Reference number: A93182; MUID:82213905; PMID:6919539
A/Contents: annotation; omega-aminocarboxylic acid binding sites
R/Valli, Z.; Pecthy, L.
J. Biol. Chem. 259, 13690-13694, 1984
A/Title: The fibrin-binding site of human plasminogen. Arginines 32 and 34 are essential
A/Reference number: A92438; MUID:55054794; PMID:6094526
A/Contents: annotation; fibrin binding site; omega-aminocarboxylic acid binding site
R/Cao, Y.; Li, R.W.; Davidson, D.; Schaller, J.; Marti, D.; Soehndel, S.; McCance, S.G.;
J. Biol. Chem. 271, 29461-29467, 1996
A/Title: Kringie domains of human angiotensin. Characterization of the anti-proliferativ
A/Reference number: A58811; MUID:97067211; PMID:8910633
A/Contents: annotation
R/Lijnen, H.R.; Ugwu, F.; Bini, A.; Collen, D.
Biochemistry 37, 4699-4702, 1998
A/Title: Generation of an angiotensin-like fragment from plasminogen by stromelysin-1 (M

A:Reference number: A58812; MUID:9548733; PMID:9548733
A:Contents: annotation
R:Tulinsky, A.; Molichek, A.M.
submitted to the Brookhaven Protein Data Bank, July 1991
A:Reference number: A51341; PDB:1PE4
A:Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 376-454
R:Tulinsky, A.; Wu, T.P.
submitted to the Brookhaven Protein Data Bank, July 1991
A:Reference number: A51488; PDB:2PK4
A:Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 375-454
R:Wu, T.P.; Tulinsky, A.
submitted to the Brookhaven Protein Data Bank, August 1993
A:Reference number: A51911; PDB:1PEK
A:Contents: annotation; X-ray crystallography, 2.48 angstroms, residues 102-181
R:Padmanabhan, K.; Tulinsky, A.
submitted to the Brookhaven Protein Data Bank, April 1994
A:Reference number: A52408; PDB:1PMK
A:Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 377-454
R:Tulinsky, A.; Mathews, I.I.
submitted to the Brookhaven Protein Data Bank, December 1995
A:Reference number: A65244; PDB:1CEA
A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
R:Tulinsky, A.; Mathews, I.I.
submitted to the Brookhaven Protein Data Bank, December 1995
A:Reference number: A65245; PDB:1CEB
A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
R:Molichek, A.M.; Tulinsky, A.; Ravichandran, K.G.
Biochemistry 30, 10576-10588, 1991
A:Title: Crystal and molecular structure of human plasminogen kringle 4 refined at 1.9 Å
A:Reference number: A58819; MUID:92031502; PMID:1657148
A:Contents: annotation
R:Wu, T.P.; Padmanabhan, K.; Tulinsky, A.; Molichek, A.M.
Biochemistry 30, 10589-10594, 1991
A:Title: The refined structure of the epsilon-aminocaproic acid complex of human plasmin
A:Reference number: A58818; MUID:92031503; PMID:1657149
A:Contents: annotation
R:de Vos, A.M.; Ultsch, M.H.; Kelley, R.F.; Padmanabhan, K.; Tulinsky, A.; Westbrook, M.
Biochemistry 31, 270-279, 1992
A:Title: Crystal structure of the kringle 2 domain of tissue plasminogen activator at 2.2 Å
A:Reference number: A39483; MUID:92118803; PMID:1310033
A:Contents: annotation; X-ray crystallography, 2.4 angstroms
R:Stec, B.; Teeter, M.M.; Whitlow, M.; Yamano, A.
submitted to the Brookhaven Protein Data Bank, June 1995
A:Reference number: A65980; PDB:1KXN
A:Contents: annotation; X-ray crystallography, 1.67 angstroms, residues 376-454
R:Rejzante, M.; Llinas, M.
submitted to the Brookhaven Protein Data Bank, August 1996
A:Reference number: A65803; PDB:1HPJ
A:Contents: annotation; conformation by (1)H-NMR, residues 103-181
R:Rejzante, M.; Llinas, M.
submitted to the Brookhaven Protein Data Bank, August 1996
A:Reference number: A65804; PDB:1HPK
A:Contents: annotation; conformation by (1)H-NMR, residues 103-181
R:Rejzante, M.R.; Llinas, M.
Biochemistry 221, 927-937, 1994
A:Title: (1)H-NMR assignments and secondary structure of human plasminogen kringle 1.
A:Reference number: A34645; MUID:94237157; PMID:8181475
A:Contents: annotation; conformation by (1)H-NMR, residues 96-184
R:Rejzante, M.R.; Llinas, M.
Biochemistry 221, 939-949, 1994
A:Title: Solution structure of the epsilon-aminohexanoic acid complex of human plasminogen
A:Reference number: A58817; MUID:94237158; PMID:8181476
A:Contents: annotation; conformation by (1)H-NMR
C:Comment: Plasminogen is synthesized by the kidney and is present in plasma and many other body fluids.
C:Comment: Plasminogen is converted to plasmin by plasminogen activators (see PIR:UKU0 and PIR:FGU0B).
C:Comment: Plasmin is inactivated by alpha-2-antiplasmin (see PIR:THU02) immediately after its release from plasminogen.
C:Comment: Plasminogen is formed by autolytic cleavage of plasmin under artificial conditions.
C:Comment: Streptolysin 1 (see PIR:KCH01) acts on plasminogen to produce angiotensin. The latter is a potent vasoconstrictor.
C:Gene: GDB:PLG
C:Genetics:

A:Cross-references: GDB:119498; OMIM:173350
A:Map position: 6q26-6q27
A:Insertions: 17/1, 62/2, 98/1, 136/2, 183/1, 223/2, 263/1, 317/2, 366/1, 419/2, 480/1, 529

C:Function:
A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a va
as the walls of the graatian follicle; also activates the urokinase-type plasminogen act
A:Pathway: fibrinolysis
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: angiotensin inhibitor; blood; duplication; fibrinolysis; glycoprotein; hydr
F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-810/Product: plasminogen #status experimental <PRO>
F:20-96/Domain: activation peptide #status experimental <ATP>
F:79-466/Product: angiotensin #status experimental <AS1>
F:97-580/Domain: plasmin chain A #status experimental <CHA>
F:97-580/Domain: plasmin chain B #status experimental <CHB>
F:103-181/Domain: kringle homology <KR1>
F:185-262/Domain: kringle homology <KR2>
F:275-352/Domain: kringle homology <KR3>
F:377-454/Domain: kringle homology <KR4>
F:481-560/Domain: kringle homology <KR5>
F:550-580,581-810/Product: microplasmin #status experimental <MMT>

Query Match 10.2%; Score 142; DB 1; Length 810;
Best Local Similarity 37.2%; Pred. No. 0.00058;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

OY 24 CFPMNGHLVREDQTSAPRLRCLNMLDA-----OSGLASAPVGAGNHSCRNPPEDPRG 78
Db 103 CKTGNGKNRHGMSTKTKNGITCKKKSSTPRRPFSPATHPSEGL-EENYCNPNDNRPG 161
OY 79 PWCYSGEAGVPEKR--PCEDLRCEP 102
Db 162 PWCYTLD---PEKRYDYCDLCEER 183

RESULT 7
A:Accession: A40522
Plasma (EC 3.4.21.7) precursor - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 28-Feb-1992 #sequence_revision 17-Apr-1993 #text_change 16-Jul-1999
C:Accession: A40522
R:Kanalas, J.J.; Makker, S.P.
J Biol Chem. 266, 10825-10829, 1991
A:Title: Identification of the rat Heymann nephritis autoantigen (GP330) as a receptor
A:Reference number: A40522; PMID:91250378; PMID:1645711
A:Accession: A40522
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-169 <AN>
A:Cross-references: GB:MG6833; NID:g206215; PIDN:AAA1884.1; PID:g9554488
A>Note: the authors translated the codon TCT for residue 76 as Ala
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:34-112,55-95,83-107/Disulfide bonds: #status predicted

Query Match 9.9%; Score 138; DB 2; Length 169;
Best Local Similarity 31.8%; Pred. No. 0.00021;
Matches 35; Conservative 13; Mismatches 48; Indels 14; Gaps 5;

OY 24 CFPMNGHLVEDQTSAPRLRCLNMLDA-----OSGLASAPVGAGNHSCRNPPEDPRG 78
Db 34 CYQNGNSYRGSTSTNTTKCCQSWMTFPHSHSKTPANPDSDL-EMNYCRNPDDORIG 92
OY 79 PWCYSGEAGVPEKR--PCEDLRCEPTTSQALPAFTTEIQASEGPGADE 126
Db 93 PWCFTTD----PSVRWEYNCLRKCSETGGV--ASAIVPVSPAPETSE 136

RESULT 8
UKBAY
u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon

C:Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)
C:/Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jun-1999
C:/Accession: S14687, S08651
R/AU, Y.P.T., Wang, T.W.; Clowes, A.W.
Nucleic Acids Res. 18, 3411, 1990
A>Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminogen
A:/Reference number: S14687; MUID:90287734; PMID:2113276
A:/Accession: S14687
A/Molecule type: mRNA
A/Residues: 1-433 <NU>
C:/Cross-references: EMBL:X51935; NID:G38130; PIDN:CA66200.1; PID:G38131
C/Superfamily: urokinase-type plasminogen activator; BGF homology; kringle homology; try
C/Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-20/Dominin: signal sequence #status predicted <SIG>
F:21-116/Product: plasminogen activator chain A #status predicted <ACH>
F:30-61/Dominin: BGF homology <BGF>
F:69-150/Dominin: kringle homology <KR>
F:168-433/Product: plasminogen activator chain B #status predicted <BCH>
F:178-421/Dominin: trypsin homology <TRY>
F:167-298, 208-224, 216-287, 315-384, 347-363, 374-402/Dissulfide bonds: #status predicted
F:123, 274, 378/Active site: His, Asp, Ser #status predicted
F:324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.9%; Score 138; DB 1; Length 433;
Best Local Similarity 32.7%; Pred. No. 0.00061;
Matches 32; Conservative 14; Mismatches 36; Indels 16; Gaps 4;

OY 24 CFWDNGHLVREDQTSPPGLRLCLNWIDA-----QGSLAPVSAGANHSYCNPDEDP 76
| :
Db 69 CYENGGHFHYGKAKSTDTMGSRCLANNWATVLQOTYHARSPALQLGLGKINYCRNDP-NR 127

OY 77 RGPWCYSGEAGVPEK-----RPCEDLRCPTTSQL 108
| :
Db 128 RRPWCYV--QVGLKKRVQECMVNCCADKKPSPEEL 163

RESULT 9
UCS061
macrophage-stimulating protein 1 precursor - rat
C/Species: Rattus norvegicus (Norway rat)
C:/Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 16-Jun-2000
C:/Accession: JCS061
R/Oshiro, K.; Iwama, A.; Matsumoto, K.; Ezaki, T.; Sakamoto, O.; Hamaguchi, I.; Takasu, I
Biochem. Biophys. Res. Commun. 227, 273-280, 1996
A>Title: Molecular cloning of rat macrophage-stimulating protein and its involvement in
A:/Reference number: JCS061; MUID:97011126; PMID:8858136
A:/Accession: JCS061
A/Molecule type: mRNA
A/Residues: 1-716 <OHS>
C/Cross-references: EMBL:X95096; NID:g1669718; PIDN:CA64473.1; PID:g1669719
C/Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C/Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
C/Keywords: duplication; glycoprotein; growth factor; kringle
F:1-31/Dominin: signal sequence #status predicted <SIG>
F:32-488, 489-716/Product: macrophage-stimulating protein 1 #status predicted <MAT>
F:32-488/Dominin: macrophage-stimulating protein 1 alpha chain #status predicted <ACH>
F:110-186/Dominin: kringle homology <KR1>
F:191-268/Dominin: kringle homology <KR12>
F:292-370/Dominin: kringle homology <KR13>
F:379-457/Dominin: kringle homology <KR14>
F:489-716/Dominin: macrophage-stimulating protein 1 beta chain #status predicted <BCH>
F:489-709/Dominin: trypsin homology <TRY>
F:72, 305, 620/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.9%; Score 137; DB 1; Length 716;
Best Local Similarity 27.3%; Pred. No. 0.0013;
Matches 48; Conservative 12; Mismatches 56; Indels 60; Gaps 8;

OY 1 FLAVQAFLYSNNMLAAFGSG-----GCFWNDGHLVREDQTSPPGLRLCLNW 48
Db 80 LLPWTQ----HSLRAQLHHSSLCDFPKQKYVFRTCLMDNGASRYGVARTADGLPCQAW 134

OY 49 ---DPAQSGLASPVSAGANHSYCNPDEDPGPMCVYS----- 84

135 SRRFPDHDKXHTPTPKNGL--EENFCGRNPDSDPRPCWCTTNTSVRSRQSGIKSCREAVCW 193
 Db 85 -----GGAQVPEK-RCCE--DLRCPEI-----TSQALPATTEIQASBGP 122
 194 CNGEDYRGVSDVETSGRECRQWMDLQHPHSHPFPEKFPDALKDNYCRNPASERP 249
 RESULT 10
 UKPG
 U-plasminogen activator (EC 3.4.21.73) precursor - pig
 N:Alternate names: uPA
 C:Species: Sus scrofa domestica (domestic pig)
 C>Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998
 C:Accession: A00932
 R:Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.
 Nucleic Acids Res. 12, 9525-9541, 1984
 A>Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.
 A:Reference number: A00932; MUID:85087954; PMID:6096832
 A:Accession: A00932
 A:Molecule type: DNA
 A:Residues: 1-240, 'H', 242-442 <NAG1>
 R:Nagamine, Y.
 A:Experimental source: kidney cell line LLC-PK1
 submitted to the Protein Sequence Database, December 1986
 A:Reference number: A37566
 A:Contents: annotation; correction to residue 241
 C:Genetics:
 A:Insertions: 19/3; 31/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F:33-64/Domain: EGF homology <EGF>
 F:72-153/Domain: kringle homology <KR>
 F:190-442/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
 F:190-430/Domain: trypsin homology <TRY>
 F:152/binding site: carboxylate (Asn) (covalent) #status predicted
 F:119-310, 220-226, 428-299, 354-393, 356-372, 383-411/Disulfide bonds: #status predicted
 F:235, 286, 378/Active site: His, Asp, Ser #status predicted
 Query Match 9.8%; Score 135.5; DB 1; Length 442;
 Best local similarity 36.9%; Pred. No. 0.001;
 Matches 38; Conservative 7; Mismatches 37; Indels 21; Gaps 5;
 QY 24 CPMFNDHLTRDQTPAPGLRCLNMLDAQSGL----ASAPVS---GAGHSYCRPPDDP 76
 Db 72 CFEQNGHSIRKGAANTVTGSRPLCPMNSATVLTNTYAHNRPDALQLGKHNCRPPD-NQ 130
 QY 77 RGFPCYVS-----GEGVP-----EKRPEDLRCPEITTSQ 106
 Db 131 RRPFCYVQVGLKQLVQECWPNRCSGSGSHRPAYDQKVPSTPE 173
 RESULT 11
 JS0599
 t-plasminogen activator (EC 3.4.21.68) beta precursor - common vampire bat
 N:Alternate names: tissue plasminogen activator
 C:Species: Desmodus rotundus (common vampire bat)
 C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
 C:Accession: JS0599
 R:Kraetzschnur, J.; Haendler, B.; Langer, G.; Boitoli, W.; Bringmann, P.; Alagon, A.; Don
 Gene 105, 229-237, 1991
 A>Title: The plasminogen activator family from the salivary gland of the vampire bat Des
 A:Reference number: JS0597; MUID:92039036; PMID:1937019
 A:Accession: JS0599
 A:Molecule type: mRNA
 A:Residues: 1-431 <KRA>
 A:Cross-References: GB:M63989; NID:g166076; PIDN:AAA31594.1; PID:g166077
 C:Superfamily: tissue plasminogen activator; BGF homology; fibrinectin type I repeat hom
 C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-36/Domain: propeptide #status predicted <PRO>

F.37-431/Product: plasminogen activator beta #status predicted <PLA>
F.41-74/Domain: EGF homology <EGF>
F.82-163/Domain: kringle homology <KR>
F.180-425/Domain: trypsin homology <TRY>
F.152-46-65, 66-74, 82-165, 103-145, 134-158, 168-299, 211-227, 215-228, 313-388/Disulfide bond
F.119-352/Binding site: carbohydrate (asn) (covalent) #status predicted
F.178-180/Cleavage site: His-Ser (plasmin) #status predicted
F.226, 275, 382/Active site: His, Asp, Ser #status predicted
F.345-361, 378-406/Disulfide bonds: #status predicted

Query Match	9.7%;	Score 134;	DB 2;	Length 431;
Best Local Similarity	38.6%;	Pred. No. 0.0013;		
Matches 27; Conservative	9;	Mismatches 20;	Indels 14;	Gaps 3;

```

QY 24 CEMDNGHLYREDQTSAPDGLRCLMUAOSGL-----ASAPVSGAGNHSYCRMPD 73
Db 82 CYKDGGVYTRGTWSTSGAQCLNW--NSNLLTRRTYNGRRSDAILTLGAGNNHRCRMPD 138
QY 74 EDDRGKWCYV 83
Db 139 NNSK-FWCYV 147

```

RESULT 12
A34369
t-Plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderma lyra)
C:Species: Megaderma lyra
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A34369
R:Gardell, S.J., Duong, L.T., Diehl, R.E., York, J.D., Hare, T.R., Register, R.B., Jacob
U. Biol. Chem. 264, 17947-17952, 1989
A:Title: Isolation, characterization, and cDNA cloning of a vampire bat salivary plasmin
A:Reference number: A34369; PMID:30036867; PMID:2509450
A:Accession: A34369
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-477 <GR>
A:Cross-references: GB:005082; NID:J166080; PIND:AAA1356.1; PID:G166081
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I
C:Keywords: fibronlysis; glycoprotein; hydrolyase; kringleg; serine proteinase
F:1-21/Domain: signal sequence #status predicted <PRO>
F:23-36/Domain: propeptide #status predicted <PRO>
F:37-477/Product: plasminogen activator #status predicted <PLA>
F:47-79/Domain: fibronectin type I repeat homology <1FA>
F:87-120/Domain: RGF homology <RGF>
F:128-209/Domain: kringleg homology <KRK>
F:226-271/Domain: cystein homology <TRY>
F:287-359/Domain: cystein homology <TRY>
F:44-76, 79-87, 96, 97-109, 111-120, 128-209, 149-191, 180-204, 214-345, 257-273, 265-334, 359-477
F:42-321, 428/Active site: His, Asp, Ser #status predicted
F:212-321, 428/Active site: His, Asp, Ser #status predicted

Query Match	9.7%;	Score 134;	DB 1;	Length 477;
Best Local Similarity	38.6%;	Pred. NO. 0.0015;		
Matches 27;	Conservative	9;	Mismatches 20;	Indels 14;
				Gaps 3;

```

QY 24 CPMIDNGHLYREDDQSSPAPGRLCLMLDAOSGL-----ASAPVSGAHNSYCRPPD 73
Db 128 CYKDGQGVYRTGVTSTESGAGCLLNM--NSNLLTRRTYGRRSAILLGLGHHNYCRPPD 164
QY 74 EDPGRGFMCYV 83
Db 185 NNSK--PMCYV 193

```

RESULT 13
 JS0598
 L-plasminogen activator (EC 3.4.21.68) alpha-2 precursor - common vampire bat
 N:Alternate names: tissue plasminogen activator
 C:Species: Desmodus rotundus (common vampire bat)
 C:Date: 31-Mar-1992 #sequence_rev:31-Mar-1992 #ext_change 16-Jul-1999
 C:Accession: JS0598
 R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Bojdel, W.; Briggmann, P.; Alagon, A.; Dorst, Gene 105, 225-237, 1991
 /Title: The plasminogen activator family from the salivary gland of the vampire bat Desm

A:Reference number: JS0597; MUID:92039036; PMID:1937019
A:Accession: JS0598
A:Molecule type: mRNA
A:Residues: 1-477 <R>
A:Cross-references: GB:ME3988; NID:g166074; PIDD:MAA1593.1; PID:g166075
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat homology
C:Keywords: fibrinolysis; glycoprotein; hydroxylase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:12-36/Domain: propeptide #status predicted <PRO>
F:37-477/Product: plasminogen activator alpha-2 #status predicted <PIA>
F:42-79/Domain: fibronectin type I repeat homology <1FA>
F:87-120/Domain: EGF homology <EGF>
F:128-209/Domain: kringle homology <KR>
F:226-471/Domain: trypsin homology <TRY>
F:442-77-70-79,87-98-59-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4
F:155,368/Binding site: carbohydrate (asn) (covalent) #status predicted
F:125-326/Cleavage site: His-Ser (plasma) #status predicted
F:127,321,428/Active site: His, Asp, Ser #status predicted

Query March	9.7%;	Score 134;	DB 2;	Length 477;
Best Local Similarity	38.6%;	Pred. No. 0.0015;		
Matches 27; Conservative	9;	Mismatches 20;	Indels 14;	Gaps 3

```

QY      24  CPMNDGHIHYEDDOTSPAGELRLNMLYDAOSGL-----ASAPVSGAGNHSYCRNPD  73
Dd      128  CYKDGGLVHYGTWSTSESAQCIIMW--NSNLTLLRTYNGRRSDAILTGLGHHNNYCRNPD  168
QY      74  EDRGRPMWCVY  83
Dd      185  NNSK-PMWCVY  193

```

RESULT 14

A40332
macrophage-stimulating protein 1 precursor - mouse

N:Alternate names: hepatocyte growth factor-like protein

C:Species: Mus musculus (house mouse)

C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 18-Jun-1999

C:Accession: A40332; B40332

R:Deegen, S.U.F.; Swartz, L.A.; Han, S.; Jamison, C.S.

Biochemistry 30, 9781-9791, 1991

A>Title: Characterization of the mouse cDNA and gene coding for a hepatocyte growth factor

A:Reference number: A40332; MUID:92002017; PMID:183957

A:Accession: A40332

A:Molecule type: DNA

A:Residues: 1-716 <DEG>

A:Cross-references: GB:M74180; NID:g193831; PIDN:AA50166.1; PID:g193832

A:Accession: B40332

A:Molecule type: mRNA

A:Residues: 1-18, 'P', 20-716 <DEG2>

A:Cross-references: GB:M74181; NID:g193833; PIDN:AA50167.1; PID:g193834

C:Gene(s):

A:Inttrons: 18/1, 67/2, 105/1, 143/2, 188/1, 229/2, 269/1, 334/2, 378/1, 412/2, 458/1, 477/1

C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor

C:Superfamily: hepatocyte growth factor; krigle homology; trypsin homology

C:Keywords: duplication; glycoprotein; growth factor; krigle

F:1/31/Domain: signal sequence #status predicted <Sig>

F:1/31/488-489-716/Product: macrophage-stimulating protein 1 #status experimental <MNT>

Query Match 9.7%; Score 134; DB 1; Length 716;
 Best Local Similarity 28.2%; Pred. No. 0.0024;
 Matches 42; Conservative 9; Mismatches 46; Indels 52; Gaps 71

QY 1 LIAWQAFIVSNMILIAEAYGSG-----GCFMDNGHLYEDQTSAPAGRLCLNW 48
||| : ||| ||| ||| ||| |||

Db 80 LLPWTQ-----HSLHTQLYHSSLCILFQKQDYVTRCTIMDNGVSRYGTAVRTAGLPCQAW 134
QY 49 ---LDAOSGLASAPVSGAGNHSYCNRPDEDPGPMCYVS----- 84
Db 135 SRPFPNDHKYTPTPKNGLT-BENFCRNPDGDRGPMCYTNRSVRFQSCGIKTCREAVCVL 193
QY 85 -----GEAGVPEK-RPCE--DLRCPEP 103
Db 194 CNGEDYRGEVDVTESGREGCCQWMDLOPHS 222

RESULT 15

JS0600
t-plasminogen activator (EC 3.4.21.68) gamma precursor - common vampire bat
M/Alternate names: tissue plasminogen activator
C/Species: Desmodus rotundus (common vampire bat)
C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C/Accession: JS0600
R/Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Dor
Gene 105, 229-237, 1991
A/TITLE: The plasminogen activator family from the salivary gland of the vampire bat Des
A/Reference number: JS0597; MUID:92039036; PMID:1937019
A/Accession: JS0600
A/Molecule type: mRNA
A/Residues: 1-194 <KRA>
A/Cross-references: GB:M63990; NID:G166078; PIDN:AAA31595.1; PID:G166079
A/Note: the authors translated the codon ATC for residue 75 as Thr
C/Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C/Keywords: fibrinolysis; glycoprotein; hydrolase; kringie; serine proteinase
F/1-21/Domain: signal sequence #status predicted <SIG>
F/122-36/Domain: propeptide #status predicted <PRO>
F/137-394/Product: plasminogen activator gamma #status predicted <PLA>
F/45-126/Domain: kringie homology <KR>
F/143-388/Domain: trypsin homology <TRY>
F/145-126,66-108,97-121,131-262,174-190,182-251,276-351,308-324,341-369/Disulfide bonds:
F/142-143/Cleavage site: His-Ser (plasmin) #status predicted
F/189,238,345/Active site: His, Asp, Ser #status predicted
F/115/Binding site: carbohydrate (Aen) (covalent) #status predicted

Query Match

9.5%; Score 132; DB 2; Length 394;

Best Local Similarity 32.7%; Pred. No. 0.0017;
Matches 33; Conservative 9; Mismatches 39; Indels 20; Gaps 5;

QY 16 AEAYS--GGCFWNGHLIREDQTSAPGLRCLNW-----LDAOSGLASAPVSGAGNH 66
Db 35 SRAVGDPRATCYKQGVYTRGTWSTESGAQCINMNSNLIRRTYNGRMPEAVKLGJGNH 94
QY 67 SYCRNPDEDPRGPWCYV-----SGEAGVPEKRPCEDLRC 100
Db 95 NYCRRPDGASK-PWCYIKARKFTSSCSVP---VCSKATC 131

Search completed: April 7, 2003, 08:43:53
Job time : 21 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 08:34:32 ; Search time 11 Seconds
(without alignments)
987.891 Million cell updates/sec

Title: US-10-057-951-2_COPY_2_263

Perfect score: 1388
Sequence: 1 LLAMVQAFIVSNMLIAEAYG.....PYVDPQEGSTPLMGQAGTPCA 262

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	160.5	11.6	566	1 TPA_BOVIN	Q28198 bos taurus
2	154.5	11.1	562	1 TPA_HUMAN	P00750 homo sapien
3	154	11.1	559	1 TPA_RAT	P19637 rattus norv
4	150	10.8	553	1 HGFA_MOUSE	Q91096 mus musculu
5	147.5	10.6	559	1 TPA_MOUSE	P11214 mus musculu
6	146.5	10.6	655	1 HGFA_HUMAN	Q04756 homo sapien
7	142	10.2	810	1 PLMN_HUMAN	P00747 homo sapien
8	138	9.9	169	1 PLMN_RAT	O01177 rattus norv
9	138	9.9	433	1 UROK_PAPCY	P16227 papio cynoc
10	135.5	9.8	442	1 UROK_PIG	P04185 sus scrofa
11	134	9.7	431	1 URTB_DESRO	P98121 desmodus ro
12	134	9.7	477	1 URT2_DESRO	P15638 desmodus ro
13	134	9.7	716	1 HGFL_MOUSE	P26928 mus musculu
14	132	9.5	394	1 URTG_DESRO	P49150 desmodus ro
15	132	9.5	431	1 UROK_HUMAN	P00749 homo sapien
16	130	9.4	810	1 PLMN_ERIEU	Q29485 etinaceus e
17	128.5	9.3	434	1 UROK_CHICK	P15120 gallus galli
18	127	9.1	433	1 UROK_BOVIN	Q05589 bos taurus
19	126.5	9.1	728	1 HGF_MOUSE	Q08048 bos taurus
20	124.5	9.0	810	1 PLMN_MACMU	P12545 macaca mula
21	124	8.9	622	1 THRB_HUMAN	P00734 homo sapien
22	124	8.9	728	1 HGF_HUMAN	P11210 homo sapien
23	124	8.9	812	1 PLMN_BOVIN	P06868 bos taurus
24	123	8.9	593	1 FA12_BOVIN	P98140 bos taurus
25	122.5	8.8	458	1 APOA_HUMAN	P08519 homo sapien
26	122	8.8	603	1 FA12_CAVPO	Q04962 cavia porce
27	121.5	8.8	728	1 HGF_RAT	P19945 rattus norv
28	120.5	8.7	333	1 PLMN_CANFA	P80009 canis faml
29	120	8.6	433	1 UROK_MOUSE	P06869 mus musculu
30	120	8.6	790	1 PLMN_PIG	P06867 sus scrofa
31	119.5	8.6	477	1 URT1_DESRO	P98119 desmodus ro
32	118.5	8.5	615	1 FA12_HUMAN	P00748 homo sapien
33	118	8.5	432	1 UROK_RAT	P29598 rattus norv

34	117	8.4	625	1 THRB_BOVIN	P00735 bos taurus
35	116	8.4	1420	1 APOA_MACMU	P14417 macaca mula
36	115.5	8.3	711	1 HGFL_HUMAN	P26927 homo sapien
37	114.5	8.2	812	1 PLMN_MOUSE	P20918 mus musculu
38	114	8.2	473	1 KREM_MOUSE	Q29943 mus musculu
39	114	8.2	473	1 KREM_RAT	Q29484 rattus norv
40	114	8.2	475	1 KREM_HUMAN	Q96mu8 homo sapien
41	114	8.2	618	1 THRB_MOUSE	P19221 mus musculu
42	109.5	7.9	1709	1 SN_HUMAN	Q9bzz2 homo sapien
43	107	7.7	617	1 THRB_RAT	P18292 rattus norv
44	103.5	7.5	343	1 PLMN_SHEEP	P1286 ovis aries
45	101.5	7.3	325	1 PLMN_PETMA	P33574 petromyzon

ALIGNMENTS

RESULT 1
ID TPA_BOVIN STANDARD; PRT; 566 AA.
AC Q28198;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tpa)
DE (t-PA) (t-plasminogen activator).
GN PLAT.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_Taxid=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Ravn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators uPA
and tPA";
RU Int. Dairy J. 5:605-617(1995).
CC -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOMEN PLASMINOGEN
TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
MANY OTHER PHYSIOLOGICAL EVENTS.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
plasminogen to form plasmin.
CC -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
BOND.
CC -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE FORM BY A CLEAVAGE AFTER
ARG-314 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XIa.
CC -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X85800; CA59795.1; --
CC HSP: P00750; 1RTP.
CC MEROPS: S01.232; --
CC InterPro: IPR001314; Chymotrypsin.
CC InterPro: IPR000561; BGF-like.
CC InterPro: IPR000083; Fibronectin.

```
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF000089; EGF; 1.
DR Pfam; PF000339; fn1; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spec; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 21
FT PROPEP 22 33
FT CHAIN 34 566
FT CHAIN 34 314
FT CHAIN 315 566
FT DOMAIN 40 82
FT DOMAIN 83 121
FT DOMAIN 128 209
FT DOMAIN 219 300
FT DOMAIN 315 566
FT ACT_SITE 361 361
FT ACT_SITE 410 410
FT ACT_SITE 517 517
FT DISULFID 42 72
FT DISULFID 70 79
FT DISULFID 87 98
FT DISULFID 92 109
FT DISULFID 111 120
FT DISULFID 128 209
FT DISULFID 149 191
FT DISULFID 180 204
FT DISULFID 219 300
FT DISULFID 240 282
FT DISULFID 271 295
FT DISULFID 303 434
FT DISULFID 346 362
FT DISULFID 354 423
FT DISULFID 448 523
FT DISULFID 480 496
FT DISULFID 513 541
FT CARBOHYD 153 153
FT CARBOHYD 487 487
SQ SEQUENCE 566 AA; 63701 MW; 2EB6BEB4E3276C3 CRC64;

Query Match 11.64; Score 160.5; DB 1; Length 566;
Best Local Similarity 35.54; Pred. No. 3.4e-06;
Matches 38; Conservative 11; Mismatches 45; Indels 13; Gaps 4;

QY 15 LAEAYSGGCGFWMDNGHLVREDQTSAPAGLRCLNWLDAQSLGASAPVS-----GAGN 65
DB 119 LCEIDATATCYKQGVAVRGFWSTAESGACAW--NSSGLAKPYSGRRPNAIRGLGN 176
QY 66 HSYCRNDDDEPRGHWCVSGEAGVPEKRPEDIRCPETTSQALPAFT 112
DB 177 HNYCRNDDQSK-FWCVF-KAGKTYSEFCSTPACAKVAEEDDCYT 221

RESULT 2
```

```
TPA HUMAN
ID TPA_HUMAN STANDARD; PRT; 562 AA.
AC P00750; Q15103;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (TPA)
DE (c-Pa) (t-plasminogen activator) (A1teplase) (Retelase).
GN PLAT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=83115262; PubMed=6337343;
RA Penica D., Holmes W.E., Kohr W.J., Harkins R.N., Vehar G.A., H.L.,
RA Ward C.A., Bennett W.F., Yelverton E., Seeburg P.H., Heyneker H.L.,
RA Coedel D.V., Collen D.;
RT "Cloning and expression of human tissue-type plasminogen activator
RT cDNA in E. coli.";
RL Nature 301:214-221 (1983).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal lung;
RX MEDLINE=86262579; PubMed=3133640;
RA Sasaki H., Saito Y., Hayashi M., Otsuka K., Niwa M.;
RT "Nucleotide sequence of the tissue-type plasminogen activator cDNA
RT from human fetal lung cells.";
RL Nucleic Acids Res. 16:5695-5695 (1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88054470; PubMed=2824147;
RA Reddy V.B., Garramone A.J., Sasak H., Wei C.-M., Watkins P., Gall J.,
RA Hsiung N.;
RT "Expression of human uterine tissue-type plasminogen activator in
RT mouse cells using BSV vectors.";
RL DNA 6:461-472 (1987).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=86196143; PubMed=3009482;
RA Friezner Degen S.J., Rajput B., Reich E.;
RT "The human tissue plasminogen activator gene.";
RL J. Biol. Chem. 261:6972-6985 (1986).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=84298137; PubMed=6089198;
RA Ny T., Elgh F., Lund B.;
RT "The structure of the human tissue-type plasminogen activator gene:
RT correlation of intron and exon structures to functional and
RT structural domains.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359 (1984).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=86284200; PubMed=3090401;
RA Harris T.J., Patel T., Marston F.A., Little S., Emtage J.S.,
RA Opendecker G., Volckaert G., Rombauts W., Billiau A., Somer P.;
RT "Cloning of cDNA coding for human tissue-type plasminogen activator
RT and its expression in Escherichia coli.";
RL Mol. Biol. Med. 3:279-292 (1986).
RN [7]
RP SEQUENCE OF 212-361 FROM N.A.
RX MEDLINE=83169656; PubMed=6572897;
RA Edlund T., Ny T., Raahy M., Heden L.-O., Palm G., Holmgren E.,
RA Josephson S.;
RT "Isolation of cDNA sequences coding for a part of human tissue
RT plasminogen activator.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:345-352 (1983).
RN [8]
RP SEQUENCE OF 1-36 FROM N.A.
RX MEDLINE=85289338; PubMed=3161893;
RA Fisher R., Waller E.K., Grossi G., Thompson D., Tizard R.,
```

RA Schleuning W.-D.;
 RT "Isolation and characterization of the human tissue-type plasminogen
 RT activator structural gene including its 5' flanking region."
 RL J. Biol. Chem. 260:11223-11230(1985).
 RN [9]
 RP SEQUENCE OF 31-562 FROM N.A.
 RX MEDLINE=91291340; PubMed=1368681;
 RA Itagaki Y., Yasuda H., Morinaga T., Mitsuda S., Higashio K.;
 RT "Purification and characterization of tissue plasminogen activator
 RT secreted by human embryonic lung diploid fibroblasts, IMR-90 cells."
 RL Agric. Biol. Chem. 55:1225-1232(1991).
 RN [10]
 RP SEQUENCE OF 36-562.
 RX TISSUE-Melanoma;
 RA MEDLINE=8500468; PubMed=6433976;
 RA Pohl G., Kaelin-Lang A., Bergsdorf N., Wallen P., Joernvall H.;
 RT "Tissue plasminogen activator: peptide analyses confirm an indirectly
 RT derived amino acid sequence, identify the active site serine residue,
 RT establish glycosylation sites, and localize variant differences."
 RL Biochemistry 23:3701-3707(1984).
 RN [11]
 RP SEQUENCE OF 33-52 AND 311-330.
 RX TISSUE-Melanoma;
 RA MEDLINE=83209620; PubMed=6682760;
 RA Wallen P., Pohl G., Bergsdorf N., Raandy M., Ny T., Joernvall H.;
 RT "Purification and characterization of a melanoma cell plasminogen
 RT activator."
 RL Eur. J. Biochem. 132:681-686(1983).
 RN [12]
 RP SEQUENCE FROM N.A. (SHORT ISOFORM).
 RX TISSUE=Umbilical vein;
 RA MEDLINE=90191219; PubMed=2107528;
 RA Siebert P.D., Fong K.;
 RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from
 RT human endothelial cells."
 RL Nucleic Acids Res. 18:1086-1086(1990).
 RN [13]
 RP STRUCTURE OF CARBOHYDRATES.
 RX MEDLINE=90092112; PubMed=2513186;
 RA Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.;
 RT "Carbohydrate structure of recombinant human uterine tissue
 RT plasminogen activator expressed in mouse epithelial cells."
 RL Eur. J. Biochem. 186:273-286(1989).
 RN [14]
 RP CARBOHYDRATE-LINKAGE SITE THR-36.
 RX MEDLINE=91159408; PubMed=1900431;
 RA Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.;
 RT "Tissue plasminogen activator has an O-linked fucose attached to
 RT threonine-61 in the epidermal growth factor domain."
 RL Biochemistry 30:2311-2314(1991).
 RN [15]
 RP DISULFIDE BONDS IN KRINGLE 2.
 RX MEDLINE=91244765; PubMed=1645336;
 RA Vilhous C.J., Wilhelm O.G., Hassell T., Jaskunas S.R., Bang N.U.;
 RT "Disulfide pairing of the recombinant kringle-2 domain of tissue
 RT plasminogen activator produced in *Bscherichia coli*."
 RL J. Biol. Chem. 266:10070-10072(1991).
 RN [16]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF CATALYTIC DOMAIN.
 RX MEDLINE=96200985; PubMed=8613982;
 RA Lamba D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,
 RA Bode W.;
 RT "Type 2.3 A crystal structure of the catalytic domain of recombinant
 RT two-chain human tissue-type plasminogen activator."
 RL J. Mol. Biol. 258:117-135(1996).
 RN [17]
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.
 RX MEDLINE=97449126; PubMed=9305622;
 RA Renatus M., Engh R.A., Stubbs W.T., Huber R., Fischer S., Kohnert U.,
 RA Bode W.;
 RT "Lysine 156 promotes the anomalous proenzyme activity of tPA: X-ray
 RT crystal structure of single-chain human tPA."
 RL EMBO J. 16:4797-4805(1997).

RN [18]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF KRINGLE 2.
 RX MEDLINE=92118803; PubMed=1310033;
 RA de Vos A., Ultsch M.H., Kelley R.F., Padmanabhan K., Tulinsky A.,
 RA Westbrook M.L., Kosiakof A.A.;
 RT "Crystal structure of the kringle 2 domain of tissue plasminogen
 RT activator at 2.4-A resolution."
 RL Biochemistry 31:270-279(1992).
 RN [19]
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=90122799; PubMed=2558718;
 RA Byeon I.-U.L., Kelley R.F., Llinas M.;
 RT "1H NMR structural characterization of a recombinant kringle 2 domain
 RT from human tissue-type plasminogen activator."
 RL Biochemistry 28:9350-9360(1989).
 RN [20]
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=91200042; PubMed=1901789;
 RA Byeon I.-U.L., Kelley R.F., Llinas M.;
 RT "Kringle-2 domain of the tissue-type plasminogen activator. 1H-NMR
 RT assignments and secondary structure."
 RL Eur. J. Biochem. 197:155-165(1991).
 RN [21]
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=92106329; PubMed=1762144;
 RA Byeon I.-U.L., Llinas M.;
 RT "Solution structure of the tissue-type plasminogen activator kringle
 RT 2 domain complexed to 6-aminocaproic acid an antifibrinolytic
 RT drug."
 RL J. Mol. Biol. 222:1035-1051(1991).
 RN [22]
 RP STRUCTURE BY NMR OF 38-85.
 RX MEDLINE=92292163; PubMed=1602484;
 RA Downing A.K., Driscoll P.C., Harvey T.S., Dudgeon T.J., Smith B.O.,
 RA Baron M., Campbell I.D.;
 RT "The solution structure and backbone dynamics of the fibronectin type
 RT I and epidermal growth factor-like pair of modules of tissue-type
 RT plasminogen activator."
 RL J. Mol. Biol. 225:821-833(1992).
 RN [23]
 RP STRUCTURE BY NMR OF 36-126.
 RX MEDLINE=96027104; PubMed=7582899;
 RA Smith B.O., Downing A.K., Driscoll P.C., Dudgeon T.J., Campbell I.D.;
 RT "The solution structure and backbone dynamics of the fibronectin type
 RT I and epidermal growth factor-like pair of modules of tissue-type
 RT plasminogen activator."
 RL Structure 3:823-833(1995).
 CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN
 CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
 CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
 CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
 CC MANY OTHER PHYSIOLOGICAL EVENTS.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: BINDS TO FIBRIN WITH HIGH AFFINITY. THIS INTERACTION

Query Match 11.1%; Score 154.5; DB 1; Length 562;
 Best Local Similarity 39.6%; Pred. No. 1.1e-05;
 Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
 QY 24 CFWDNGLHYREDQTSAPDLRLCLNMLDAOSGLASAPVS-----GAGNSHYCRNPDE 74
 DB 127 CYDDGIGSYRGWTGSAEAGAECTNW--NSSALAQKPYSGRRPDAIRLGLGNHNYCRNPD 184
 QY 75 DPRGPCVYSGEAGVPEKRPCEDLRCPETTS 105
 DB 185 DSK-PWCYVF-KAKGKSSSEFCSTPACSEGN 213

RESULT 3
 TPA RAT
 ID TPA RAT
 AC P19637;
 STANDARD; PRT; 559 AA.

01-FEB-1991 (Rel. 17, Created)
01-NOV-1991 (Rel. 20, Last sequence update)
15-JUN-2002 (Rel. 41, Last annotation update)
Tissue-type plasminogen activator precursor (SC 3.4.21.68) (CPA)
(t-PA) (t-plasminogen activator).
PLAT.
Rattus norvegicus (Rat).
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_taxid=10116;
[1]
SEQUENCE FROM N.A.
MEDLINE=89170114; PubMed=318445;
Ny T.; Leonardsson G.; Huseh A.J.W.;
"Cloning and characterization of a cDNA for rat tissue-type
plasminogen activator";
DNA 7:671-677(1988).
[2]
SEQUENCE FROM N.A.
MEDLINE=9010448; PubMed=2105315;
Feng P.; Ohlsson M.; Ny T.;
The structure of the tPA-less rat tissue-type plasminogen activator
gene. Species-specific sequence divergences in the promoter predict
differences in regulation of gene expression.";
J. Biol. Chem. 265:2022-2027(1990).
-1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN
TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CONTROLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
MANY OTHER PHYSIOLOGICAL EVENTS.
-1- CATALYTIC ACTIVITY: Specific cleavage of Arg|-Val bond in
plasminogen to form plasmin
-1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
BOND.
-1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
-1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
-1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
-1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
-1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE 1 DOMAIN.
-1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.ebi.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; M23697; AAA41812.1; -;
EMBL; M31197; AAA42261.1; -;
EMBL; M31185; AAA42261.1; JOINED.
EMBL; M31186; AAA42261.1; JOINED.
EMBL; M31187; AAA42261.1; JOINED.
EMBL; M31188; AAA42261.1; JOINED.
EMBL; M31189; AAA42261.1; JOINED.
EMBL; M31190; AAA42261.1; JOINED.
EMBL; M31191; AAA42261.1; JOINED.
EMBL; M31192; AAA42261.1; JOINED.
EMBL; M31193; AAA42261.1; JOINED.
EMBL; M31194; AAA42261.1; JOINED.
EMBL; M31195; AAA42261.1; JOINED.
EMBL; M31196; AAA42261.1; JOINED.
EMBL; A19618; CA01482.1; -;
PIR; A31597; A31597.
PIR; A35029; A35029.
HSSP; P00750; 1RTE.
MEROPS; S01.232; -;
InterPro; IPR001314; Chymotrypsin.

InterPro; IPR000561; EGF-like.
InterPro; IPR000083; Fibrinctn.
InterPro; IPR000001; Kringle.
InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; fn1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_2; 1.
DR PROSITE; PS01186; EGF_1; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_2; 2.
DR PROSITE; PS00070; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 17
FT PROPEP 18 29
FT CHAIN 30 559
FT CHAIN 30 308
FT CHAIN 309 559
FT DOMAIN 36 78
FT DOMAIN 79 117
FT DOMAIN 124 205
FT DOMAIN 213 294
FT ACT SITE 309 559
FT ACT SITE 355 355
FT ACT SITE 404 404
FT ACT SITE 510 510
FT DISULFD 38 68
FT DISULFD 66 75
FT DISULFD 83 94
FT DISULFD 88 105
FT DISULFD 107 116
FT DISULFD 124 205
FT DISULFD 145 187
FT DISULFD 176 200
FT DISULFD 213 294
FT DISULFD 234 276
FT DISULFD 265 289
FT DISULFD 297 428
FT DISULFD 340 356
FT DISULFD 348 417
FT DISULFD 442 516
FT DISULFD 474 490
FT DISULFD 506 534
FT CARBOHYD 149 149
FT CARBOHYD 481 481
FT CONFLICT 380 380
SQ SEQUENCE 559 AA; 62903 MW; 7DBD3809CD1DC921 CRC64;
Query Match 11.1%; Score 154; DB 1; Length 559;
Best Local Similarity 32.4%; Pred. No. 1.2e-05;
Matches 45; Conservative 14; Mismatches 54; Indels 26; Gaps 6;
QY 24 CPWNGHLYREDOTSPAPGLRCNMLDAGSLAAPS-----GAGNHSYGRNDE 74
DB 124 CREGGQITRTGTSTANGABECTNM--NSSALSOXPSIARRPNAIKLGANHYCRNPDR 181
QY 75 DFRGWCYVSGEAGVPEKRPCEDLRCP-----TTSQALPAFTTEIQEASDEGPG 123
DB 182 DVK-FWCYVF-KAGKYTEFCSTPACPKGPTEDCYVGKGVYTGRTSHFTT--SKASCLPFW 237

QY 124 ADEVQFANALPARESEA 142
Db 238 NSMILIGKTYTAMRANSQA 256

RESULT 4
HGFA_MOUSE STANDARD; PRT; 653 AA.
AC 09R038: 09JRV4; 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (BC 3.4.21.-) (HGF activator) (HGFA).
GN HGFA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Itoh H., Kataoka H., Koono H.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RA van Aelbeberg J.S., Sehgal S., Kukes A., Brady C., Barasch J., Yang J., Huan Y.;
RT "Activation of HGF by endogenous HGF activator is required for mesangial kidney morphogenesis in vitro."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY CONVERTING IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM (BY SIMILARITY).
CC -I- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -I- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -I- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -I- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE II DOMAIN.
CC -I- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -I- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by, and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF099017; AAF02489.1; -
DR EMBL: AF224724; AAF34712.1; -
DR HSSP: P00763; IDPO.
DR MEROPS: S01.228; -
DR MGP: MGI.185281; Hgfac.
DR InterPro: IPR001314; Cytomotrypsin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF 2.
DR InterPro: IPR000562; FN Type II.
DR InterPro: IPR000083; Fibrinctn1.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_Protease_Try.
DR Pfam: PF00008; EGF; 2.
DR Pfam: PF00038; fn1; 1.
DR Pfam: PF00040; fn2; 1.
DR Pfam: PF00051; kringle; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PRO0722; CHYMOTRYPSIN.
DR PRINTS: PRO0013; FNTYPEII.

DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle; 1.
DR ProDom: PD000995; FN Type II; 1.
DR SMART: SM00181; EGF_2.
DR SMART: SM0058; FN1; 1.
DR SMART: SM0059; FN2; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; TRYD_SPC; 1.
DR PROSITE: PS00022; EGF 1; 2.
DR PROSITE: PS01186; EGF 2; 1.
DR PROSITE: PS01253; FIBRONECTIN_1; 1.
DR PROSITE: PS00023; FIBRONECTIN_2; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
DR PROSITE: PS050240; TRYPsin_DOM; 1.
DR PROSITE: PS00134; TRYPsin_HIS; 1.
DR PROSITE: PS00135; TRYPsin_SER; 1.
KM Hydroxylase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
KW EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 29
FT PROPEP 30 369
FT CHAIN 370 405
FT FT CHAIN 406 653
FT FT DOMAIN 105 145
FT FT DOMAIN 157 195
FT FT DOMAIN 197 237
FT FT DOMAIN 238 276
FT FT DOMAIN 283 364
FT FT DOMAIN 406 653
FT FT ACT SITE 445 445
FT FT ACT SITE 495 495
FT FT ACT SITE 596 596
FT FT DISULFID 105 130
FT FT DISULFID 119 145
FT FT DISULFID 161 172
FT FT DISULFID 166 183
FT FT DISULFID 185 194
FT FT DISULFID 199 227
FT FT DISULFID 225 234
FT FT DISULFID 242 253
FT FT DISULFID 247 264
FT FT DISULFID 266 275
FT FT DISULFID 283 346
FT FT DISULFID 304 346
FT FT DISULFID 335 359
FT FT DISULFID 392 519
FT FT DISULFID 430 446
FT FT DISULFID 438 508
FT FT DISULFID 533 602
FT FT DISULFID 565 581
FT FT DISULFID 592 620
FT FT CARBOHYD 39 39
FT FT CARBOHYD 47 47
FT FT CARBOHYD 63 63
FT FT CARBOHYD 287 287
FT FT CARBOHYD 466 466
FT FT CARBOHYD 544 544
FT FT CONFLICT 164 164
SQ SEQUENCE 653 AA; 70567 MW; 88B4B2055DF7DC CRC64;

Query Match 10.8%; Score 150; DB 1; Length 653;
Best Local Similarity 32.6%; Pred. No. 3; 1e-05;
Matches 46; Conservative 17; Mismatches 46; Indels 32; Gaps 7;

QY 24 CPWDNGHYREDQSPAPGRCJLNM-----LDQSGLASPVSGAGNHSYCRNPDD 75
Db 283 CFLGNGTEYRGVASTAASGLSLMNSDLYOELHYDS-VAAVVLGLSPHAYCNPDXD 341
QY 76 PRGPMCVVSGEAGVPEKPCEDLRCPEFTSQALPAFTTEIQEASBSPGADDEVQFAPANA 135
Db 342 ER-PWCYVVKDNLMSWE-----YCRLTACESLARVHSQTFE-----ILA--A 380

QY 136 LPARSEAAVCPVIGISQVR 156
DB 381 LP--ESAPAVRPTCGKRRKR 399

RESULT 5
TPA_MOUSE STANDARD; PRT; 559 AA.

AC P1214;
DT 01-JUN-1989 (Rel. 11, Created)
DT 01-JUN-1989 (Rel. 11, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator).
GN PLAT.

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8808703; PubMed=2826484;
RA Rickles R.J., Darrow A.L., Strickland S.;
RT "Molecular cloning of complementary DNA to mouse tissue plasminogen
RT activator mRNA and its expression during F9 teratocarcinoma cell
RT differentiation.";
RT J. Biol. Chem. 263:1563-1569(1988).
RL -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN
CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC MANY OTHER PHYSIOLOGICAL EVENTS.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC BOND.
CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: J03520; AAA0470.1; -.
DR PIR: A29941; A29941.
DR HSSP: P00750; IASH.
DR MEROPS: S01.232; -.
DR MGD: MGI:97610; PLAT.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000083; Fibrinctn.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00008; EGF; 1.
DR Pfam: PF00039; fn1; 1.
DR Pfam: PF00051; kringle; 2.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle; 2.

DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KM Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 17 PROBABLE.
FT PROPEP 18 29
FT CHAIN 30 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 30 308 TISSUE-TYPE PLASMINOGEN ACTIVATOR A
FT CHAIN 309 559 CHAIN.
FT CHAIN 309 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR B
FT DOMAIN 36 78 FIBRONECTIN TYPE-1.
FT DOMAIN 79 117 EGF-LIKE.
FT DOMAIN 124 205 KRINGLE 1.
FT DOMAIN 213 294 KRINGLE 2.
FT DOMAIN 309 559 SERINE PROTEASE.
FT ACT SITE 355 355 CHARGE RELAY SYSTEM.
FT ACT SITE 404 404 CHARGE RELAY SYSTEM.
FT ACT SITE 510 510 CHARGE RELAY SYSTEM.
FT DISULFID 38 68 BY SIMILARITY.
FT DISULFID 66 75 BY SIMILARITY.
FT DISULFID 83 94 BY SIMILARITY.
FT DISULFID 88 105 BY SIMILARITY.
FT DISULFID 107 116 BY SIMILARITY.
FT DISULFID 124 205 BY SIMILARITY.
FT DISULFID 145 187 BY SIMILARITY.
FT DISULFID 176 200 BY SIMILARITY.
FT DISULFID 213 294 BY SIMILARITY.
FT DISULFID 234 276 BY SIMILARITY.
FT DISULFID 265 289 BY SIMILARITY.
FT DISULFID 297 428 INTERCHAIN (BY SIMILARITY).
FT DISULFID 340 356 BY SIMILARITY.
FT DISULFID 348 417 BY SIMILARITY.
FT DISULFID 442 516 BY SIMILARITY.
FT DISULFID 474 490 BY SIMILARITY.
FT DISULFID 506 534 BY SIMILARITY.
FT CARBOHYD 149 149 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 559 AA; 63110 MW; 4ACEB57DCA282A5 CRC64;

Query Match 10.6%; Score 147.5; DB 1; Length 559;
Best Local Similarity 37.0%; Pred. No. 4, 3e-05;
Matches 34; Conservative 11; Mismatches 34; Indels 13; Gaps 4;

QY 24 CPWDNGHYRSDQSPAPGLRCNWLDAOSGLASAVS-----GAGNHSYGRNDE 74
DB 124 CREEGGITYRGTSWASGACLIW--NSSVLKPKYNNRPAIKLGLGNHNYCRNPD 181

QY 75 DRRGWCYVSGEAGVPEKRPCEDLRCPEITTSQ 106
DB 182 DLK-PWCYVF-KAGKYTFEFCSTPACPKGSE 211

RESULT 6
HGFA_HUMAN STANDARD; PRT; 655 AA.
AC Q04756; Q14726;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF
DE activator) (HGFA).
GN HGFAc.

OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RC TISSUE=Liver, and Serum;
 RX MEDLINE=93252878; PubMed=7683665;
 RA Miyazawa K., Shimomura T., Kitamura A., Kondo J., Morimoto Y.,
 RA Kitamura N.;
 RT "Molecular cloning and sequence analysis of the cDNA for a human
 RT serine protease responsible for activation of hepatocyte growth
 RT factor. Structural similarity of the protease precursor to blood
 RT coagulation factor XII.";
 RL J. Biol. Chem. 268:10024-10028(1993).
 RN [2]
 RP SEQUENCE OF 40-655 FROM N.A.
 RA Zhao S., Odell C.;
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 CC -1 FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY
 CC CONVERTING IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM.
 CC -1 SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A
 CC DISULFIDE BOND.
 CC -1 SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN
 CC PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM.
 CC -1 TISSUE SPECIFICITY: LIVER.
 CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1 SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
 CC -1 SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
 CC -1 SIMILARITY: CONTAINS 1 KRINGEIN TYPE II DOMAIN.
 CC -1 SIMILARITY: CONTAINS 1 KRINGEIN DOMAIN.
 CC -1 CAUTION: IT IS UNCERTAIN WHETHER MET-1 IS THE INITIATOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D14012; BA003113.1; -;
 DR EMBL; Z69923; CA93803.1; -;
 DR PIR; A46688; A46688.
 DR HSSP; P00763; IDPO.
 DR MEROPS; S01.228; -;
 DR Genew; HGNC:4894; HGFAC.
 DR MIM; 604552; -;
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR000742; EGF 2.
 DR InterPro; IPR000562; FN_Type_II.
 DR InterPro; IPR000083; Fibrinctn.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00039; fn1; 1.
 DR Pfam; PF00040; fn2; 1.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00013; FNYPEII.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 1.
 DR ProDom; PD000995; FN_Type_II; 1.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00058; FN1; 1.
 DR SMART; SM00059; FN2; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.

DR PROSITE; PS00023; FIBRONECTIN 2; 1.
 DR PROSITE; PS00021; KRINGLE 1; 1.
 DR PROSITE; PS00070; KRINGLE 2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolyase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
 KW EGF-like domain; Repeat; Zymogen.
 FT SIGNAL 1 30
 FT PROPEP 31 372
 FT CHAIN 373 407
 FT CHAIN 408 655
 FT DOMAIN 108 148
 FT DOMAIN 160 198
 FT DOMAIN 200 240
 FT DOMAIN 241 279
 FT DOMAIN 286 367
 FT DOMAIN 408 655
 FT ACT_SITE 447 447
 FT ACT_SITE 497 497
 FT ACT_SITE 538 598
 FT ACT_SITE 598 598
 FT DISULFID 108 133
 FT DISULFID 122 148
 FT DISULFID 164 175
 FT DISULFID 169 186
 FT DISULFID 188 197
 FT DISULFID 202 230
 FT DISULFID 228 237
 FT DISULFID 245 256
 FT DISULFID 250 267
 FT DISULFID 269 278
 FT DISULFID 286 367
 FT DISULFID 307 349
 FT DISULFID 338 362
 FT DISULFID 394 521
 FT DISULFID 432 448
 FT DISULFID 440 510
 FT DISULFID 535 604
 FT DISULFID 567 583
 FT DISULFID 594 622
 FT CARBOHYD 48 48
 FT CARBOHYD 290 290
 FT CARBOHYD 468 468
 FT CARBOHYD 492 492
 FT CARBOHYD 546 546
 FT CONFLICT 644 644
 SQ SEQUENCE 655 AA; 70681 MW; 2CF72F1B862ED7 CRC64;
 Query Match 10.6%; Score 146.5; DB 1; Length 655;
 Best Local Similarity 36.9%; Pred. No. 6.2e-05;
 Matches 41; Conservative 10; Mismatches 41; Indels 19; Gaps 6;
 QY 24 CFMNGHLYREDDQTSPPRGRCINM-----LDAGSLASPVGAGNHSYCRPPDED 75
 DB 286 CFLNGGTGYGVASTSAGSLCLAMNSDLYOELHYDS-VGAALALGLPHACRCPPDND 344
 QY 76 PRGFWCVSGEAGVP-----EKRPCEDLRCPEITSQALPATTETIOE-ASEG 121
 DB 345 ER-FWCTVVDASLSWEXCLNACESL-----TRVQSPDLATLPBPASP 390
 RESULT 7
 ID PLIN HUMAN STANDARD; PRT; 810 AA.
 AC P00747;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
 GN Plg.
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90202879; PubMed=2318848;
RA Petersen T.E., Martzen M.R., Ichinose A., Davie E.W.;
RT "Characterization of the gene for human plasminogen, a key proenzyme
in the fibrinolytic system.";
RL J. Biol. Chem. 265:6104-6111(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87162490; PubMed=3030813;
RA Forsgren M., Raden B., Israelsson M., Larsson K., Heden L.-O.;
RT "Molecular cloning and characterization of a full-length cDNA clone
for human plasminogen.";
RL FEBS Lett. 213:254-260(1987).
RN [3]
RP SEQUENCE OF 20-810.
RA Sottrup-Jensen L., Petersen T.E., Magnusson S.;
RL Submitted (JUL-1977) to the EMBL data bank.
RN [4]
RP SEQUENCE OF 292-810 FROM N.A.
RX MEDLINE=85023311; PubMed=6148961;
RA Malinowski D.P., Sadler J.E., Davie E.W.;
RT "Characterization of a complementary deoxyribonucleic acid coding for
human and bovine plasminogen.";
RL Biochemistry 23:4243-4250(1984).
RN [5]
RP SEQUENCE OF 20-100.
RX MEDLINE=75093329; PubMed=122932;
RA Wiman B., Wallen P.;
RT "Structural relationship between 'glutamic acid' and 'lysine' forms
of human plasminogen and their interaction with the NH₂-terminal
activation peptide as studied by affinity chromatography.";
RL Eur. J. Biochem. 50:489-494(1975).
RN [6]
RP SEQUENCE OF 95-580; 581-626; 657-700 AND 732-810.
RA Sottrup-Jensen L., Claeys H., Zajdel M., Petersen T.E., Magnusson S.;
RL (in) Davidson J.F., Roman R.M., Samama M.M., Desnoyers P.C. (eds.);
RT Progress in Chemical Fibrinolysis and Thrombolysis, pp.3:191-209,
RL Raven Press, New York (1978).
RN [7]
RP SEQUENCE OF 483-604.
RX MEDLINE=76043692; PubMed=126863;
RA Wiman B., Wallen P.;
RT "Amino-acid sequence of the cyanogen-bromide fragment from human
plasminogen that forms the linkage between the plasmin chains.";
RL Eur. J. Biochem. 58:539-547(1975).
RN [8]
RP SEQUENCE OF 581-810.
RX MEDLINE=77252245; PubMed=142009;
RA Wiman B.;
RT "Primary structure of the B-chain of human plasmin.";
RL Eur. J. Biochem. 76:129-137(1977).
RN [9]
RP ACTIVE SITE.
RX MEDLINE=73149248; PubMed=4694729;
RA Robbins K.C., Bernabe P., Arzadon L., Summaria L.;
RT "The primary structure of human plasminogen. II. The histidine loop
of human plasmin: light (B) chain active center histidine sequence.";
RL J. Biol. Chem. 248:1631-1633(1973).
RN [10]
RP ACTIVE SITE.
RX MEDLINE=69234739; PubMed=4240117;
RA Groskopf W.R., Summaria L., Robbins K.C.;
RT "Studies on the active center of human plasmin. Partial amino acid
sequence of a peptide containing the active center serine residue.";
RL J. Biol. Chem. 244:3590-3597(1969).
RN [11]
RP OMEGA-AMINOCARBOXYLIC ACID-BINDING SITES.
RX MEDLINE=82213905; PubMed=6919539;
RA Trexler M., Vail Z., Patchy L.;

RT "Structure of the omega-aminocarboxylic acid-binding sites of human
plasminogen. Arginine 70 and aspartic acid 56 are essential for
RT binding of ligand by kringle 4.";
RL J. Biol. Chem. 257:7401-7406(1982).
RN [12]
RP FIBRIN AND OMEGA-AMINOCARBOXYLIC ACID BINDING SITES.
RX MEDLINE=85054794; PubMed=6094526;
RA Vail Z., Patchy L.;
RT "The fibrin-binding site of human plasminogen. Arginines 32 and 34
are essential for fibrin affinity of the kringle 1 domain.";
RL J. Biol. Chem. 259:13690-13694(1984).
RN [13]
RP PHOSPHORYLATION SITE SER-597.
RX MEDLINE=97345939; PubMed=9201958;
RA Wang H., Protok M., Bretthauer R.K., Castellino F.J.;
RT "Serine-578 is a major phosphorylation locus in human plasma
plasminogen.";
RL Biochemistry 36:8100-8106(1997).
RN [14]
RP CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=88185329; PubMed=3356193;
RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamberling J.P.,
RA Gerwig G.J., Van Halbeek H., Vliegenthart J.F.;
RT "The N- and O-linked carbohydrate chains of human, bovine and porcine
plasminogen. Species specificity in relation to sialylation and
RT fucosylation patterns.";
RL Eur. J. Biochem. 173:57-63(1988).
RN [15]
RP CARBOHYDRATE-LINKAGE SITE SER-268.
RX MEDLINE=97207306; PubMed=9054441;
RA Pizzo-Shepherd S.R., Stevens R.D., Andon N.L., Enghild J.J.,
RA Pizzo S.V.;
RT "Evidence for a novel O-linked sialylated triaccharide on Ser-248 of
RT human plasminogen 2.";
RL J. Biol. Chem. 272:7408-7411(1997).
RN [16]
RP CHARACTERIZATION OF ANGIOTATIN, AND PARTIAL SEQUENCE.
RX MEDLINE=95042728; PubMed=7525077;
RA O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,
RA Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
RT "Angiotatin: a novel angiogenesis inhibitor that mediates the
RT suppression of metastases by a Lewis lung carcinoma.";
RL Cell 79:315-328(1994).
RN [17]
RP CHARACTERIZATION OF ANGIOTATIN.
RX MEDLINE=97238710; PubMed=9102221;
RA Sim B.K., O'Reilly M.S., Liang H., Fortier A.H., He W., Madsen J.W.,
RA Lapcevysh R., Nacy C.A.;
RT "A recombinant human angiotatin protein inhibits experimental primary
RT and metastatic cancer.";
RL Cancer Res. 57:1329-1334(1997).
RN [18]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 374-461.
RX MEDLINE=92031502; PubMed=1657148;
RA Mulichak A.M., Tulinsky A., Ravichandran K.G.;
RT "Crystal and molecular structure of human plasminogen kringle 4
RT refined at 1.9-A resolution.";
RL Biochemistry 30:10576-10588(1991).
RN [19]
RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 374-461.
RX MEDLINE=92031503; PubMed=1657149;
RA Wu T.-P., Padmanabhan K., Tulinsky A., Mulichak A.M.;
RT "The refined structure of the epsilon-aminocaproic acid complex of
RT human plasminogen kringle 4.";
RL Biochemistry 30:10589-10594(1991).
RN [20]
RP X-RAY CRYSTALLOGRAPHY (1.67 ANGSTROMS) OF 376-454.
RX Stec B., Yamano A., Whitlow M., Teeter M.M.;
RT "Structure of human plasminogen kringle 4 at 1.68 Angstrom and 277 K.
RT A possible structural role of disordered residues.";
RL Acta Crystallogr. D 53:169-178(1997).
RN [21]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 102-181.


```

RX MEDLINE=96180681; Pubmed=8611560;
RA Mathews I.I., Vanderhoff-Hanover P., Castellino F.J., Tulinsky A.;
RT "Crystal structures of the recombinant kringle 1 domain of human
RT plasminogen in complexes with the ligands epsilon-aminocaproic acid
RT and trans-4-(aminomethyl)cyclohexane-1-carboxylic Acid.";
RL Biochemistry 35:2567-2576(1996).
RN
RP X-RAY CRYSTALLOGRAPHY (1.66 ANGSTROMS) OF 480-563.
RX MEDLINE=98198034; Pubmed=9521645;
RA Chang Y., Mochalkin I., McCance S.G., Cheng B., Tulinsky A.,
RA Castellino F.J.;
RT "Structure and ligand binding determinants of the recombinant kringle
RT 5 domain of human plasminogen.";
RL Biochemistry 37:3258-3271(1998).
RN
RP STRUCTURE BY NMR OF 96-184.
RX MEDLINE=94237157; Pubmed=8181475;
RA Rejzante M.R., Llinas M.;
RT "1H-NMR assignments and secondary structure of human plasminogen
RT kringle 1.";
RL Eur. J. Biochem. 221:927-937(1994).
RN
RP STRUCTURE BY NMR OF 96-184.
RX MEDLINE=94237158; Pubmed=8181476;
RA Rejzante M.R., Llinas M.;
RT "Solution structure of the epsilon-aminohexanoic acid complex of
RT human plasminogen kringle 1.";
RL Eur. J. Biochem. 221:939-949(1994).
RN
RP STRUCTURE BY NMR OF 183-354.
RX MEDLINE=96194155; Pubmed=8652577;
RA Seehndel S., Hu C.-K., Marti D., Affolter M., Schaller J., Llinas M.,
RA Rickli E.E.;
RT "Recombinant gene expression and 1H NMR characteristics of the
RT kringle (2 + 3) supermodule: spectroscopic/functional individuality
RT of plasminogen kringle domains.";
RL Biochemistry 35:2357-2364(1996).
RN
RP STRUCTURE BY NMR OF 374-461.
RX MEDLINE=90219023; Pubmed=2157850;
RA Atkinson R.A., Williams R.U.P.;
RT "Solution structure of the kringle 4 domain from human plasminogen by
RT 1H nuclear magnetic resonance spectroscopy and distance geometry.";
RL J. Mol. Biol. 212:541-552(1990).
RN
RP VARIANTS PHE-374 AND THR-620.

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=91250378; Pubmed=1645711;
RA Kanalas J.J., Makher S.P.;
RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a
RT receptor site for plasminogen.";
RL J. Biol. Chem. 266:10825-10829(1991).
RN
RP FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
RP A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
RP EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
RP AND INFLAMMATION. IN OVULATION IT WEAKENS THE WALLS OF THE
RP GRAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
RP ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
RP AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
RP LAMININ AND VON WILLEBRAND FACTOR.
RP CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
RP higher selectivity than trypsin. Converts fibrin into soluble
RP products.
RP -I- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
RP ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
RP FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
RP -I- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
RP IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
RP -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
RP -I- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M62832; AAA41884.1; -.
DR PIR: A40522; A40522.
DR HSP: P00747; 1PMK.
DR MEROPS: S01.233; -.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF000051; Kringle; 2.
DR ProDom: PD000395; Kringle; 2.
DR SMART: SM00130; KR; 1.
DR PROSITE: PS00021; KRINGLE 1; 1.
DR PROSITE: PS00070; KRINGLE 2; 2.
DR PROSITE: PS00240; TRYPSIN_DOM; PARTIAL.
DR PROSITE: PS00134; TRYPSIN_HIS; PARTIAL.
DR PROSITE: PS00135; TRYPSIN_SER; PARTIAL.
KW Hydrolyase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Repeat.
FT NON_TER 1 1
FT DOMAIN 1 10 KRINGLE 3 (BY SIMILARITY).
FT DOMAIN 34 112 KRINGLE 4 (BY SIMILARITY).
FT DOMAIN 139 >169 KRINGLE 5 (BY SIMILARITY).
FT DISULFID 34 112 BY SIMILARITY.
FT DISULFID 55 95 BY SIMILARITY.
FT DISULFID 83 107 BY SIMILARITY.
FT NON_TER 169 169
SQ SEQUENCE 169 AA; 18401 MW; 77A54214C49D010C CRC64;

```

Query Match 9.9%; Score 138; DB 1; Length 169;
 Best Local Similarity 31.8%; Pred. No. 7.2e-05;
 Matches 35; Conservative 13; Mismatches 48; Indels 14; Gaps 5;

```

QY 24 CFWNDGHLVREDQTSAPARGLCLMWLDA-----QSGLASAPVSGAGNHSYCRNPDEPRG 78
DB 34 CYGNGSKSYRGTSITTKKCCQSWMTPHSHSKTPANFPDSGL-ENMYCRNPDPDQRG 92
QY 79 PWCYSGEAGVPEKR--PCEDLRCPETTSQALPAPFTTEIGASBEGRGDE 126
DB 93 PWCFTTD---PSVMEYCNLRKSETGGV--AESALVPQVSPAPGRSE 136

```

RESULT 9
 UROK_PAPCY STANDARD: PRT; 433 AA.
 AC P16227;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 GN PLAU.
 OS Papio cynocephalus (Yellow baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Papio.
 OC NCBI_TaxId=9556;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thoracic aorta;
 RX MEDLINE=90287734; PubMed=2113276;
 RA Au Y.P.T., Wang T.W., Clowes A.W.;
 RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type
 RL Nucleic Acids Res. 18:3411-3411(1990).
 CC -I- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -I- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
 CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
 CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
 CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
 CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -I- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X51935; CAA36200.1; -.
 DR PIR; S14687; UKBAY.
 DR HSSP; P00749; ILMW.
 DR MEROPS; S01.231; -.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser-protease_Try.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR ProDom; PD000395; kringle; 1.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; FALSE NEG.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Plasminogen activator; Hydrolase; Serine protease; Glycoprotein;
 KW Kringle; EGF-like domain; Zymogen; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 FT CHAIN 21 176 CHAIN A (BY SIMILARITY).
 FT CHAIN 155 176 SHORT A CHAIN (A1) (BY SIMILARITY).
 FT CHAIN 178 433 CHAIN B (BY SIMILARITY).
 FT DOMAIN 26 62 EGF-LIKE.
 FT DOMAIN 69 150 KRINGLE.

FT DOMAIN 151 177 CONNECTING PEPTIDE.
 FT DOMAIN 178 433 SERINE PROTEASE.
 FT DISULFID 30 38 BY SIMILARITY.
 FT DISULFID 32 50 BY SIMILARITY.
 FT DISULFID 52 61 BY SIMILARITY.
 FT DISULFID 167 298 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 208 224 BY SIMILARITY.
 FT DISULFID 216 287 BY SIMILARITY.
 FT DISULFID 315 384 BY SIMILARITY.
 FT DISULFID 347 363 BY SIMILARITY.
 FT DISULFID 374 402 BY SIMILARITY.
 FT ACT SITE 223 223 CHARGE RELAY SYSTEM.
 FT ACT SITE 274 274 CHARGE RELAY SYSTEM.
 FT ACT SITE 378 378 CHARGE RELAY SYSTEM.
 FT CARBOHYD 324 324 N-LINKED (GLCNAC. .) (BY SIMILARITY).
 SQ SEQUENCE 433 AA; 48595 MW; 816D220FEDC8792 CRC64;
 Query Match 9.9%; Score 138; DB 1; Length 433;
 Best Local Similarity 32.7%; Pred. No. 0.00021;
 Matches 32; Conservative 14; Mismatches 36; Indels 16; Gaps 4;
 QY 24 CFWDNGHLYREDQSPAPGIRCLNWIDA-----QSLASAPVSGAGHSYCRNPDEDP 76
 DB 69 CYENGCHFYRGKASTDTMGKSCLAWSATVLOQTYHAHRSDALQLGLGHNYCRNP NR 127
 QY 77 RGPWCYVSGEAGVPER-----RPCEDLRCPERTSQAL 108
 DB 128 RRPWCYV--OVGLKQVQECMVHNCADGKKPSPPEEL 163
 RESULT 10
 UROK_PIG STANDARD: PRT; 442 AA.
 AC P04185;
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 GN PLAU.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 OC NCBI_TaxId=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=85087954; PubMed=6096832;
 RA Nagamine Y., Pearson D., Altus M.S., Reich E.;
 RT "cDNA and gene nucleotide sequence of porcine plasminogen activator.";
 RL Nucleic Acids Res. 12:9525-9541(1984).
 RN [2]
 RP REVISION TO 241.
 RA Nagamine Y.;
 RL Submitted (DEC-1986) to the PIR data bank.
 CC -I- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -I- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X01648; CAA25806.1; -.
 DR EMBL; X02724; CAA26511.1; -.
 DR PIR; A00932; UKPG.
 DR HSSP; P00749; IKDU.

DR MEROPS: S01.231. -
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam: PF00051; Kringle; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR ProDom: PD000395; Kringle; 1.
 DR SMART: SM00130; KR; 1.
 DR SMART: SM00020; TRYP_SPC; 1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; FALSE_NEG.
 DR PROSITE: PS00021; KRINGLE_1; 1.
 DR PROSITE: PS00070; KRINGLE_2; 1.
 DR PROSITE: PS50240; TRYPsin_DOM; 1.
 DR PROSITE: PS00134; TRYPsin_HIS; 1.
 DR PROSITE: PS00135; TRYPsin_SER; 1.
 DR Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
 KM Kringle; EGF-like domain; Zymogen; Signal.
 FT SIGNAL 1
 FT CHAIN 21 442 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 FT CHAIN 21 188 CHAIN A (BY SIMILARITY).
 FT CHAIN 190 442 CHAIN B (BY SIMILARITY).
 FT DOMAIN 29 65 EGF-LIKE.
 FT DOMAIN 72 153 KRINGLE.
 FT DOMAIN 154 189 CONNECTING PEPTIDE.
 FT DOMAIN 190 442 SERINE PROTEASE.
 FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .).
 FT DISULFID 33 41 BY SIMILARITY.
 FT DISULFID 35 53 BY SIMILARITY.
 FT DISULFID 55 64 BY SIMILARITY.
 FT DISULFID 179 310 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 220 236 BY SIMILARITY.
 FT DISULFID 228 299 BY SIMILARITY.
 FT DISULFID 324 393 BY SIMILARITY.
 FT DISULFID 356 372 BY SIMILARITY.
 FT DISULFID 383 411 BY SIMILARITY.
 FT ACT_SITE 235 235 CHARGE RELAY SYSTEM.
 FT ACT_SITE 286 286 CHARGE RELAY SYSTEM.
 FT ACT_SITE 387 387 CHARGE RELAY SYSTEM.
 FT ACT_SITE 241 241 Q -> H (IN REF. 1; CAA25806).
 FT CONFLICT 242 242 O -> H (IN REF. 1; CAA25511).
 FT CONFLICT 288 288 A -> GS (IN REF. 1; CAA25806).
 FT CONFLICT 288 288 A -> GS (IN REF. 1; CAA25806).
 FT SEQUENCE 442 AA; 49116 MW; EB32FCE501321EB CRC64;

Query_Match 9.8%; Score 135.5; DB 1; Length 442;
 Best Local Similarity 36.9%; Pred. No. 0.00034;
 Matches 38; Conservative 7; Mismatches 37; Indels 21; Gaps 5;

Qy 24 CFWDNGHLVREDQTPAPGLRCINWDAOSGL---ASAPVS---GAGNHSYCRNPDED 76
 Db 72 CFEQGHSHYRGKANTNGGRPCLPMSATVLLNTYAHHPDLAQGLGKHNVCNPD-NQ 130
 Qy 77 RGPACVYS-----GEAGVP-----EKRPCEDLRCPEPTSQ 106
 Db 131 RRPACVYGVGKQVQGVQECWPCFNGSGESHPRAYDGNKPFSTE 173

RESULT 11
 URB DESRO STANDARD; PRT; 431 AA.
 AC P98121;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Salivary plasminogen activator beta precursor (EC 3.4.21.68) (DSPA beta).
 OS Desmodus rotundus (Vampire bat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
 OC Desmodontinae; Desmodus.
 NCBI_taxid=9430;

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=salivary gland;
 RX MEDLINE=92039036; PubMed=1937019;
 RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
 RA Alagon A., Donner P., Schleuning W.D.,
 RT "The plasminogen activator family from the salivary gland of the
 RT vampire bat Desmodus rotundus: cloning and expression.";
 RL Gene 105:229-237(1991).
 RN [2]
 RP CHARACTERIZATION
 RX MEDLINE=93393059; PubMed=1309059;
 RA Schleuning W.D., Alagon A., Boidol W., Bringmann P., Petri T.,
 RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Wilt W.,
 RA Donner P.,
 RT "Plasminogen activators from the saliva of Desmodus rotundus (common
 RT vampire bat): unique fibrin specificity.";
 RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
 CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
 CC EXCLUSIVELY HAEMOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
 CC AGENT.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M63989; AAA31594.1; -
 DR HSSP: P98119; IAS1.
 DR MEROPS: S01.239; -
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR00561; EGF-like.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR001254; Ser_Protease_Try.
 DR Pfam: PF00089; EGF; 1.
 DR Pfam: PF00051; Kringle; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00018; KRINGLE.
 DR ProDom: PD000395; Kringle; 1.
 DR SMART: SM00181; EGF; 1.
 DR SMART: SM00130; KR; 1.
 DR SMART: SM00020; TRYP_SPC; 1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS00021; KRINGLE_1; 1.
 DR PROSITE: PS00070; KRINGLE_2; 1.
 DR PROSITE: PS50240; TRYPsin_DOM; 1.
 DR PROSITE: PS00134; TRYPsin_HIS; 1.
 DR PROSITE: PS00135; TRYPsin_SER; 1.
 DR Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
 KM Kringle; EGF-like domain; Signal; Multigene family.
 FT SIGNAL 1
 FT CHAIN 37 431 SALIVARY PLASMINOGEN ACTIVATOR BETA.
 FT DOMAIN 37 75 EGF-LIKE.
 FT DOMAIN 82 163 KRINGLE.
 FT DOMAIN 179 431 SERINE PROTEASE.
 FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 275 275 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 382 382 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 41 32 BY SIMILARITY.
 FT DISULFID 46 63 BY SIMILARITY.
 FT DISULFID 65 74 BY SIMILARITY.
 FT DISULFID 82 163 BY SIMILARITY.

```

FT DISULFID 103 145 BY SIMILARITY.
FT DISULFID 134 158 BY SIMILARITY.
FT DISULFID 168 299 BY SIMILARITY.
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 288 BY SIMILARITY.
FT DISULFID 313 388 BY SIMILARITY.
FT DISULFID 345 361 BY SIMILARITY.
FT DISULFID 378 406 BY SIMILARITY.
FT CARBOHYD 139 139 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 431 AA; 48221 MW; 699B5E675B162CBF CRC64;

Query Match 9.7%; Score 134; DB 1; Length 431;
Best Local Similarity 38.6%; Pred. No. 0.00045;
Matches 27; Conservative 9; Mismatches 20; Indels 14; Gaps 3;

QY 24 CFMDNCHLYREDOQTSPAPGRCJLWMDAOSGL-----ASAQVSGAGNHSYCRNP 73
DB 82 CYDQGVYRGYRWTSSESQAQCLIM--NSNLTTRTYGRSRDAITLGLGNHNYCRNP 138
QY 74 EDRGPGWCYV 83
DB 139 NNSK-PMCYV 147

RESULT 12
UR22_DESRO STANDARD; PRT; 477 AA.
ID P15638;
AC 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE Saliary plasminogen activator alpha 2 precursor (EC 3.4.21.66) (DSFA
DE alpha-2) (BAR-PA) (1-plasminogen activator).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
NCBI_Taxid=9430;
RN 11;
RP SEQUENCE FROM N.A.
RC TISSUE=saliary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar U., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the saliary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RT Gene 105:229-237(1991).
RN 12;
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=saliary gland;
RX MEDLINE=90036867; PubMed=2509450;
RA Gardell S.J., Duong L.T., Diehl R.E., York J.D., Hare T.R.,
RA Register R.B., Jacobs J.W., Dixon R.A.F., Friedman P.A.;
RT "Isolation, characterization, and cDNA cloning of a vampire bat
RT saliary plasminogen activator.";
RT J Biol. Chem. 264:17947-17952(1989).
RN 13;
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetzschmar U., Haendler B., Langer G., Baldus B., Witt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RT Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
CC AGENT.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- ENZYME REGULATION: ACTIVITY TOWARD PLASMINOGEN IS STIMULATED IN
CC THE PRESENCE OF FIBRIN I.

CC -1- SUBUNIT: MONOMER.
CC -1- DOMAIN: THE FIBRONECTIN TYPE-1 DOMAIN MEDIATES BINDING TO FIBRIN,
CC AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED
CC STIMULATION OF ACTIVITY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to licences@isb-sib.ch).
CC EMBL; M63988; AAA31593.1; -
CC EMBL; J05082; AAA31596.1; -
CC PIR; A34369; A34369.
CC HSRF; P98119; IAS1.
CC MEROPS; S01.232; -.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR000083; Fibronectin.
CC InterPro; IPR000001; Kringle.
CC pfam; PF00008; EGF_1.
CC pfam; PF00039; fn1; 1.
CC pfam; PF00051; kringle; 1.
CC pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00181; EGF_1.
CC SMART; SM00058; FN1; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00020; TRYPSIN; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01166; EGF_2; 1.
CC PROSITE; PS01253; FIBRONECTIN_1; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS50070; KRINGLE_2; 1.
CC PROSITE; PS50240; TRYPsin_DOM; 1.
CC PROSITE; PS00134; TRYPsin_HIS; 1.
CC PROSITE; PS00135; TRYPsin_SER; 1.
CC KMW plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
CC Kringle; EGF-like domain; Signal; Multigene family.
CC FT SIGNAL 1 36
CC FT CHAIN 37 477
CC FT DOMAIN 40 82
CC FT DOMAIN 83 121
CC FT DOMAIN 128 209
CC FT DOMAIN 225 477
CC FT ACT_SITE 272 272
CC FT ACT_SITE 321 321
CC FT ACT_SITE 428 428
CC FT DISULFID 42 72
CC FT DISULFID 70 79
CC FT DISULFID 87 98
CC FT DISULFID 92 109
CC FT DISULFID 111 120
CC FT DISULFID 128 209
CC FT DISULFID 149 191
CC FT DISULFID 180 204
CC FT DISULFID 214 345
CC FT DISULFID 237 273
CC FT DISULFID 265 334
CC FT DISULFID 359 434
CC FT DISULFID 391 407
CC FT DISULFID 424 452
CC FT CARBOHYD 185 185
CC FT CARBOHYD 398 398

```

FT CONFLICT 403 403 N -> K (IN REF. 2).
 FT CONFLICT 417 417 Y -> H (IN REF. 2).
 FT CONFLICT 435 435 M -> R (IN REF. 2).
 SQ SEQUENCE 477 AA; 53719 MW; 17486555COE5077C CRC64;
 Query Match 9.7%; Score 134; DB 1; Length 477;
 Best Local Similarity 38.6%; Pred. No. 0.0005;
 Matches 27; Conservative 9; Mismatches 20; Indels 14; Gaps 3;
 OY 24 CFMNGHLYREDQTSFAPGLRCLNMLDAOSGL-----ASAPVSGAGNHSYCRND 73
 DB 128 CYKQGVTVYRGWSTSGAQCIW---NSNLTFRRTYNGRRSDAITGLGNHNYCRND 184
 OY 74 EDRPGWCYV 83
 DB 185 NNSK-PWCYV 193
 RESULT 13
 HGFL_MOUSE STANDARD; PRT; 716 AA.
 ID P26928;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hepatocyte growth factor-like protein precursor (Macrophage
 DE stimulatory protein) (MSP).
 GN MST1 OR HGFL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OK NCBI_TaxID=10990;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/c; TISSUE=Liver;
 RA Fritzenr Degen S.U.; Stuart L.A.; Han S.; Jamison C.S.;
 RT "Characterization of the mouse cDNA and gene coding for a hepatocyte
 RT growth factor-like protein: expression during development."
 RL Biochemistry 30:9781-9791(1991).
 CC -1- FUNCTION: PROBABLY HAS NO PROTEOLYTIC ACTIVITY, SINCE CRUCIAL AA
 CC CHARACTERISTIC OF SERINE PROTEASES CATALYTIC SITES ARE NOT
 CC CONSERVED.
 CC -1- TISSUE SPECIFICITY: LIVER. LOWER LEVELS IN LUNG, PLACENTA AND
 CC ADRENAL.
 CC -1- DEVELOPMENTAL STAGE: IS EXPRESSED AT LOW LEVELS DURING GESTATION.
 CC JUST BEFORE BIRTH THE LEVEL INCREASES DRAMATICALLY AND REMAINS
 CC STABLE AFTERWARDS.
 CC -1- PMT: MAY BE CLEAVED AFTER AA 488, TO YIELD A TWO-CHAIN MOLECULE
 CC HEUD TOGETHER BY DISULFIDE BONDS, OR CLEAVED INTO TWO SEPARATE
 CC POLYPEPTIDES.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; M74180; AAA50166.1; -;
 DR EMBL; M74181; AAA50167.1; -;
 DR HSSP; P00747; 1KEN.
 DR MEROPS; S01.975; -;
 DR MGD; MGI:96080; HGFL.
 DR InterPro; IPR001334; Chymotrypsin.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR003014; PAN.
 DR InterPro; IPR003609; Pan_app.
 DR InterPro; IPR001254; Ser_protease TRY.
 DR Pfam; PF00024; PAN; 1.

DR Pfam; PR00051; kringle; 4.
 DR Pfam; PR00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR PRODOM; PD000395; Kringle; 4.
 DR SMART; SM00130; KR; 4.
 DR SMART; SM00473; PAN_AP; 1.
 DR SMART; SM00020; TRYD_SPC; 1.
 DR PROSITE; PS00021; KRINGLE_1; 4.
 DR PROSITE; PS00070; KRINGLE_2; 4.
 DR PROSITE; PS0240; TRYPSIN_DOM; 1.
 DR KMW; KMW; Glycoprotein; Serine protease homolog; Repeat; Signal.
 FT KRINGLE; 1
 FT CHAIN 19 716
 FT DOMAIN 19 109
 FT DOMAIN 110 186
 FT DOMAIN 191 268
 FT DOMAIN 292 370
 FT DOMAIN 379 457
 FT DOMAIN 489 716
 FT DISULFID 56 78
 FT DISULFID 60 66
 FT DISULFID 110 186
 FT DISULFID 131 169
 FT DISULFID 157 181
 FT DISULFID 191 268
 FT DISULFID 194 333
 FT DISULFID 212 251
 FT DISULFID 240 263
 FT DISULFID 292 370
 FT DISULFID 313 352
 FT DISULFID 341 364
 FT DISULFID 379 457
 FT DISULFID 400 440
 FT DISULFID 428 452
 FT DISULFID 477 593
 FT DISULFID 512 528
 FT DISULFID 607 672
 FT DISULFID 637 651
 FT DISULFID 662 690
 FT CARBOHYD 72 72
 FT CARBOHYD 173 173
 FT CARBOHYD 305 305
 FT CARBOHYD 620 620
 FT CONFLICT 19 19
 SQ SEQUENCE 716 AA; 80588 MW; BBCE02F85213ACC CRC64;
 Query Match 9.7%; Score 134; DB 1; Length 716;
 Best Local Similarity 28.2%; Pred. No. 0.00079;
 Matches 42; Conservative 9; Mismatches 46; Indels 52; Gaps 7;
 OY 1 LLAVYQAFVLSNMLLAAYGSG-----GCPMDNGHLYREDQTSFAPGLRCLNM 48
 DB 80 LHPYTO-----HSHTQGLHSLSCHLFQKQDYVRCIMDNQVSRGVARTAGLGCQAM 134
 OY 49 ---LDAOSGLASAPVSGAGNHSYCRNDEDPGPPWCYV----- 84
 DB 135 SRPFPNDHKTTPYKNG-ESNFCRNDGDPGPPWCYITNNSVRFQSGIKTCEAVCYL 193
 OY 85 -----GEAGVPEK-RPCE--DIRCPET 103
 DB 194 CNGEDYRGVDTVESGRCRCORWDLQHPHS 222
 RESULT 14
 URTG_DESRO STANDARD; PRT; 394 AA.
 ID URTG_DESRO
 AC P49150;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Salivary plasminogen activator gamma precursor (BC 3.4.21.68) (DSPA
 DE gamma).

OS Desmodus rotundus (Vampire bat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
 OC Desmodontinae; Desmodus.
 OC NCBI_TaxID=9430;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Salivary gland;
 RX MEDLINE=92039056; PubMed=1937019;
 RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
 RT "The plasminogen activator family from the salivary gland of the
 RT vampire bat Desmodus rotundus: cloning and expression.";
 RL Gene 105:229-237 (1991).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=93393059; PubMed=1309059;
 RA Schleming W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
 RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Wilt W.,
 RA Donner P.;
 RT "Plasminogen activators from the saliva of Desmodus rotundus (common
 RT vampire bat): unique fibrin specificity.";
 RL Ann. N.Y. Acad. Sci. 667:395-403 (1992).
 CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
 CC EXCLUSIVELY HAEMOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
 CC AGENT.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M63990; AAA31595.1; -;
 DR HSSP; P98119; IASI.
 DR MEROPS; S01.239; -;
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser.protease_Try.
 DR Pfam; PF000651; Kringle_1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; TRYPSIN_1.
 DR ProDom; PD000395; Kringle_1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; TRYPSIN_SPEC; 1.
 DR PROSITE; PS00021; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
 KW Kringle; Signal; Multigene family.
 FT SIGNAL 1 36
 FT CHAIN 1 394
 FT DOMAIN 45 126
 FT ACT_SITE 189 189
 FT ACT_SITE 238 238
 FT ACT_SITE 345 345
 FT DISULFID 45 126
 FT DISULFID 66 108
 FT DISULFID 97 121
 FT DISULFID 131 262
 FT DISULFID 174 190
 FT DISULFID * 182 251

FT DISULFID 276 351 BY SIMILARITY.
 FT DISULFID 308 324 BY SIMILARITY.
 FT DISULFID 341 369 BY SIMILARITY.
 FT CARBOHYD 315 315 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 394 AA; 44105 MW; 9CDD6F52F3D81FDC CRC64;
 Query Match 9.5%; Score 132; DB 1; Length 394;
 Best Local Similarity 32.7%; Pred. No. 0.0006;
 Matches 33; Conservative 9; Mismatches 39; Indels 20; Gaps 5;
 QY 16 AAYGSG--GGCWNDDGHLYRDOTSPAPGLRLTNW-----LDAGSGLASAPVSGAGNH 66
 DB 35 SKAIGPAPATCKKDDGVTRKQWTSISESAQCINWNSULIRRYNGRMPYAVKLGKGNH 94
 QY 67 SYCRNPDDEDPGRCVCY-----SGEAGYPERKPCEDLR 100
 DB 95 NCRNPDGASK-FWCYVTKARKFTSSCSVP--VCSKATC 131
 RESULT 15
 ID UROK_HUMAN STANDARD: PRT; 431 AA.
 AC P00749; Q15844; Q16618; Q969W6;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 20-MAR-1987 (Rel. 04, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 GN PLAT
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85215647; PubMed=2987867;
 RA Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Blasi F.;
 RT "The human urokinase-plasminogen activator gene and its promoter.";
 RL Nucleic Acids Res. 13:2759-2771 (1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Holmes W.E., Pennica D., Blader M., Rey M.W., Guenzler W.A.,
 RA Steffens G.J., Heynaker H.L.;
 RT "Cloning and expression of the gene for pro-urokinase in Escherichia
 RT coli.";
 RL Biotechnology 3:923-929 (1985).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86056954; PubMed=2415429;
 RA Nagai M., Hiramatsu R., Kaneda T., Hayasuke N., Arimura H.,
 RA Nishida M., Suyama T.;
 RT "Molecular cloning of cDNA coding for human prepro-urokinase.";
 RL Gene 56:183-188 (1985).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85203359; PubMed=2888571;
 RA Jacobs P., Cravador A., Lorlau R., Brockly F., Colau B., Chuchana P.,
 RA van Elsen A., Hertzog A., Bollen A.;
 RT "Molecular cloning, sequencing, and expression in Escherichia coli of
 RT human prepro-urokinase cDNA.";
 RL DNA 4:139-146 (1985).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
 RA Nickerson D.A.;
 RT Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RA Strausberg R.;
 RT Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 66-431 FROM N.A.

RX MEDLINE=84272706; PubMed=6589620;
 RA Verde P., Scoppelli M.P., Galeffi P., di Nocera P., Biasi F.;
 RT "Identification and primary sequence of an unspliced human urokinase
 poly(A)+ RNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).
 RN [8]
 RP SEQUENCE OF 21-177
 RA MEDLINE=83055084; PubMed=6754569;
 RX Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E.,
 RA Flohe L.;
 RT "The primary structure of high molecular mass urokinase from human
 urine. The complete amino acid sequence of the A chain.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).
 RN [9]
 RP SEQUENCE OF 156-176 AND 179-224.
 RX MEDLINE=83003608; PubMed=6749491;
 RA Schaller J., Nick H., Rickli E.E., Gillissen D., Lergler W.,
 RA Studer R.O.;
 RT "Human low-molecular-weight urinary urokinase. Partial
 RT characterization and preliminary sequence data of the two polypeptide
 chains.";
 RL Eur. J. Biochem. 125:251-257(1982).
 RN [10]
 RP SEQUENCE OF 158-410.
 RX MEDLINE=83055099; PubMed=6754572;
 RA Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L.;
 RT "The complete amino acid sequence of low molecular mass urokinase
 from human urine.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).
 RN [11]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE=96000858; PubMed=8591045;
 RA Spraggon C.M., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,
 RA Dobson C.M., Stuart D.I., Jones E.Y.;
 RT "The crystal structure of the catalytic domain of human
 RT urokinase-type plasminogen activator.";
 RL Structure 3:681-691(1995).
 RN [12]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.
 RX MEDLINE=20266327; PubMed=10805774;
 RA Sperl S., Jacob U., Arroyo de Prada N., Stuerbecher J., Wilhelm O.G.,
 RA Boe W., Magdolen V., Huber R., Moroder L.;
 RT "(4-antimethyl-1)phenylguanidine derivatives as nonpeptidic highly
 RT selective inhibitors of human urokinase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).
 RN [13]
 RP STRUCTURE BY NMR.
 RX MEDLINE=89127526; PubMed=2536903;
 RA Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.;
 RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-
 RT dimensional NMR.";
 RL Nature 337:579-582(1989).
 RN [14]
 RP STRUCTURE BY NMR OF 67-155.
 RX MEDLINE=93003110; PubMed=1327118;
 RA Li X., Smith R.A.G., Dobson C.M.;
 RT "Sequential 1H NMR assignments and secondary structure of the kringle
 RT domain from urokinase.";
 RL Biochemistry 31:9562-9571(1992).
 RN [15]
 RP STRUCTURE BY NMR OF 67-155.
 RX MEDLINE=94149701; PubMed=8107091;
 RA Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.;
 RT "Solution structure of the kringle domain from urokinase-type
 RT plasminogen activator.";
 RL J. Mol. Biol. 235:1548-1559(1994).
 RN [16]
 RP VARIANT LEU-141.
 RX MEDLINE=96186279; PubMed=8652631;
 RA Yoshimoto M., Uehiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K.,
 RA Sasaki Y., Hanada K.;
 RT "Characterization of single chain urokinase-type plasminogen
 RT activator with a novel amino-acid substitution in the kringle

RT structure.";
 RL Biochim. Biophys. Acta 1293:83-89(1996).
 RN [17]
 RP VARIANT LEU-141.
 RX MEDLINE=97218551; PubMed=9065988;
 RA Come B., Berczy M., Belin D.;
 RT "Detection of polymorphisms in the human urokinase-type plasminogen
 RT activator gene.";
 RL Thromb. Haemost. 77:434-435(1997).
 RN [18]
 RP ERRATUM.
 RA Come B., Berczy M., Belin D.;
 RL Thromb. Haemost. 76:973-973(1997).
 RN [19]
 RP VARIANT LEU-141.
 RX MEDLINE=97337920; PubMed=9194591;
 RA Turkmen B., Schmitt M., Schmalfeldt B., Trommler P., Hell W.,
 RA Creutzburg S., Graeff H., Magdolen V.;
 RT "Mutation analysis of the genes encoding urokinase-type plasminogen
 RT activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer.";
 RL Electrophoresis 18:686-689(1997).
 CC -1- FUNCTION: POTENT PLASMINOGEN ACTIVATOR AND IS CLINICALLY USED FOR
 CC THERAPY OF THROMBOLYTIC DISORDERS.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
 CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
 CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
 CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN.
 CC -1- PHARMACEUTICAL: Available under the name Abbotkinase (Abbott). Used
 CC in Pulmonary Embolism (PE) to initiate fibrinolysis.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X02419; CAA26268.1; -
 DR EMBL; M15476; AAA61253.1; -
 DR EMBL; D00244; BAA00175.1; -
 DR EMBL; D11143; BAA01919.1; -
 DR EMBL; X02760; CAA26535.1; -
 DR EMBL; AF377330; AAK53822.1; -
 DR EMBL; BC013575; AAH13575.1; -
 DR EMBL; K03226; AAC97138.1; -
 DR EMBL; K02286; AAA61252.1; -
 DR EMBL; A21571; CAA01559.1; -
 DR EMBL; A18397; CAA01390.1; -
 DR PIR; A00931; UKHU
 DR PIR; A32974; A32974.
 DR PDB; 1KDU; 31-OCT-93.
 DR PDB; 1LWV; 29-JAN-96.
 DR PDB; 1URK; 08-MAY-95.
 DR PDB; 1EUN; 17-MAY-00.
 DR MEROPS; S01.231; -
 DR GLYCOBASE; P00749; -
 DR Genew; HGNC; 9052; PLAU.
 DR MIM; 191840; -
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser protease_Try.
 DR Pfam; PF000051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 1.

DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYP_SPC; 1.

Query Match 9.5%; Score 132; DB 1; Length 431;
Best Local Similarity 32.7%; Pred. No. 0.00066;
Matches 33; Conservative 13; Mismatches 33; Indels 22; Gaps 5;

QY 24 CFWDNGHLYREDQTSAPAGLRCLNWLDA-----OSGLASAPVSGAGNHSYCRNPDEDP 76
Db 70 CREGNGHFTYRGKASTDTMGRRPCLPMSATVLTQTYHAHRSDDLQGLGKHNVCRNPD-NR 128
QY 77 RGPWCYVSGEAGVPEKRP-----CEDLRCPETTSQAL 108
Db 129 RRPWCYV--QVGL--KPLVQECWVHDCADGKKSSPPEEL 164

Search completed: April 7, 2003, 08:42:48
Job time : 13 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 08:39:12; Search time 33 Seconds

(without alignments)
1635,889 Million cell updates/sec

Title: US-10-057-951-2_COPY_2_263

Sequence: 1 LLAWQAFIVSNMLLAAYG.....PVDPQGSTPLMGQATGCA 262

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SPTREMBL_21.*

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_todent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvius:*
16: sp_bacteriag:*
17: sp_cheap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1388	100.0	263	Q96FE7	Q96FE7 homo sapien
2	1385	99.8	263	Q00318	Q00318 homo sapien
3	167	12.0	562	Q8S023	Q8S023 sus scrofa
4	154.5	11.1	516	Q9BU99	Q9BU99 homo sapien
5	150	10.8	653	Q8VCS4	Q8VCS4 mus musculu
6	147.5	10.6	559	Q91VP2	Q91VP2 mus musculu
7	142	10.2	810	Q15146	Q15146 mus sapien
8	139.5	10.1	385	Q25101	Q25101 hermania m
9	137	9.9	716	P70521	P70521 ratius norv
10	136.5	9.8	420	Q90504	Q90504 eptactetus
11	135	9.7	812	Q9RCW3	Q9RCW3 ratius norv
12	134.5	9.7	395	Q9BZM1	Q9BZM1 homo sapien
13	134.5	9.7	704	Q90865	Q90865 gallus galli
14	134	9.7	716	Q91XG8	Q91XG8 mus musculu
15	131.5	9.5	313	Q9PU78	Q9PU78 crocodylus
16	131	9.4	154	Q96SE8	Q96SE8 homo sapien

17	131	9.4	608	13	Q9PTW7	Q9PTW7 struthio ca
18	130.5	9.4	717	13	P70006	P70006 xenopus lae
19	129	9.3	616	6	Q97507	Q97507 sus scrofa
20	128.5	9.3	806	6	Q18783	Q18783 macropus eu
21	127	9.1	157	6	Q9TVW8	Q9TVW8 bos taurus
22	126	9.1	716	13	Q91691	Q91691 xenopus lae
23	125.5	9.0	560	4	Q14520	Q14520 homo sapien
24	123	8.9	728	6	Q9BHD9	Q9BHD9 felis silve
25	121.5	8.8	420	4	Q96GL8	Q96GL8 homo sapien
26	121.5	8.8	420	4	Q9BTP9	Q9BTP9 homo sapien
27	119	8.6	334	6	Q46507	Q46507 papio hamad
28	119	8.6	710	13	Q91402	Q91402 xenopus. he
29	117.5	8.5	812	11	Q91WJ5	Q91WJ5 mus musculu
30	114.5	8.2	113	4	Q9UR7	Q9UR7 mus sapien
31	114	8.2	113	4	Q9UR5	Q9UR5 mus sapien
32	114	8.2	202	13	Q90675	Q90675 gallus galli
33	114	8.2	452	13	Q90Y90	Q90Y90 xenopus lae
34	114	8.2	607	13	Q91001	Q91001 gallus galli
35	112.5	8.1	359	6	Q8WNR1	Q8WNR1 canis faml
36	112	8.1	567	4	Q13208	Q13208 homo sapien
37	112	8.1	594	5	P91823	P91823 caenorhabd
38	112	8.1	1145	5	Q9BKL8	Q9BKL8 alysia cal
39	111.5	8.0	597	11	Q35727	Q35727 mus musculu
40	111	8.0	726	13	Q90978	Q90978 gallus galli
41	110.5	8.0	685	5	Q24488	Q24488 drosophila
42	110.5	8.0	2358	16	Q911V8	Q911V8 streptomyc
43	108	7.8	215	13	Q42341	Q42341 gallus galli
44	106	7.6	378	13	Q90WPD	Q90WPD trachemys s
45	105.5	7.6	709	13	Q90ZNE	Q90ZNE brachydanto

ALIGNMENTS

RESULT 1

Q96FE7; PRELIMINARY; PRT; 263 AA.

AC Q96FE7; 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Unknown (protein for MGC:17330).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DDA databases.
DR EMBL; BC011049; AAH1049.1; -
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle_1.
DR ProDom; PD000395; Kringle_1.
DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_1.
DR PROSITE; PS00070; KRINGLE_2; 1.
SQ SEQUENCE 263 AA; 28234 MW; 197C3EE888FA242 CRC64;

Query Match 100.0%; Score 1388; DB 4; Length 263;
Best Local Similarity 100.0%; Pred. No. 2.6e-117;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	LLAWQAFIVSNMLLAAYGSGGCFWNGHLYREDQTSPPAGLRCLNMLDAOSGLASAPV	60
DB	2	LLAWQAFIVSNMLLAAYGSGGCFWNGHLYREDQTSPPAGLRCLNMLDAOSGLASAPV	61
QY	61	SGAGNSYCRNDEDPDPPMCYSGEAGVPEKPCEDLRCPETTSOALPAFTTEIOEASE	120
DB	62	SGAGNSYCRNDEDPDPPMCYSGEAGVPEKPCEDLRCPETTSOALPAFTTEIOEASE	121
QY	121	GGGADEVQVFAPNALPARRSRAAVQPVYIGISQVRMNSKEKDLGTVYVIGTMMYTI	180

Db 122 GPGADEVGFAPANALPARSEAAAVQPVIGISQVRMNSKEKKDGLTGYVLGITMMVII 181
 QY 181 IAIAGGIIILGYSYKRGDKLEQHDQKVCEREMORITLPLSAFTNPLCEIVDEKTVVHTS 240
 Db 182 IAIAGGIIILGYSYKRGDKLEQHDQKVCEREMORITLPLSAFTNPLCEIVDEKTVVHTS 241
 QY 241 QTPVDPQEGSTPLMGQAGTPGA 262
 Db 242 QTPVDPQEGSTPLMGQAGTPGA 263

RESULT 2

000318 PRELIMINARY; PRT; 263 AA.
 AC 000318.
 DT 01-JUN-1997 (TREMBlrel. 04, Created)
 DT 01-JUN-1997 (TREMBlrel. 04, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE WUGSC:DJ515N1.2 protein.
 GN WUGSC:DJ515N1.2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Du Z., Scheet P., Harper M.;
 RT "The sequence of H. sapiens PAC clone Rp3-515N1."
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 RN (2)
 RP SEQUENCE FROM N.A.
 RA Waterston R.
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC002073; AAB54054.1; -.
 DR HSSP; P00749; 1KDQ.
 DR InterPro; IPR000001; Kringle.
 DR Pfam; PF00051; Kringle_1.
 DR PRINTS; PR00018; KRINGLE; FALSE_NEG.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00130; KR; 1.
 DR PROSITE; PS00021; KRINGLE_2; 1.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 SQ SEQUENCE 263 AA; 28248 MW; 197C3EE8B54242 CRC64;

Query Match 99.8%; Score 1365; DB 4; Length 263;
 Best Local Similarity 99.6%; Pred. No. 4.9e-117;
 Matches 261; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIAWQAFLVSNMLLAAYGSGCFMNGHLVREDQTSAPAGRCILNMLDAQSGIASAPV 60
 Db 2 LIAWQAFLVSNMLLAAYGSGCFMNGHLVREDQTSAPAGRCILNMLDAQSGIASAPV 61
 QY 61 SGAGNHSYCRNPDEDPGRGFWCVSGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIOEASE 120
 Db 62 SGAGNHSYCRNPDEDPGRGFWCVSGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIOEASE 121
 QY 121 GPGADEVGFAPANALPARSEAAAVQPVIGISQVRMNSKEKKDGLTGYVLGITMMVII 180
 Db 122 GPGADEVGFAPANALPARSEAAAVQPVIGISQVRMNSKEKKDGLTGYVLGITMMVII 181
 QY 181 IAIAGGIIILGYSYKRGDKLEQHDQKVCEREMORITLPLSAFTNPLCEIVDEKTVVHTS 240
 Db 182 IAIAGGIIILGYSYKRGDKLEQHDQKVCEREMORITLPLSAFTNPLCEIVDEKTVVHTS 241
 QY 241 QTPVDPQEGSTPLMGQAGTPGA 262
 Db 242 QTPVDPQEGSTPLMGQAGTPGA 263

RESULT 3

08SQ23 PRELIMINARY; PRT; 562 AA.
 ID 08SQ23.
 AC 08SQ23;

DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE T-plasminogen activator.
 OS Sue scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sue.
 OC NCBI_TaxID=9823;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=EMFAMEL ORGAN;
 RA Ding Y., Xue J., Bartlett J.D.;
 RT "T-plasminogen activator in tooth tissues."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF364605; AAM00297.1; -.
 SQ SEQUENCE 562 AA; 63668 MW; F9E6B4C77CB101E8 CRC64;

Query Match 12.0%; Score 167; DB 6; Length 562;
 Best Local Similarity 36.6%; Pred. No. 1.2e-06;
 Matches 41; Conservative 12; Mismatches 33; Indels 26; Gaps 6;

QY 24 CFMNGHLVREDQTSAPAGRCILNMLDAQSGIASAPV-----GAGNHSYCRNPDE 74
 Db 127 CYEDGITYRGTWSTYSSGAEVCNM--NTSGIASMPYIGRRPDAVKGLGHNHCRPDK 184
 QY 75 DPRGFWCV--SGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIOEASGPGAD 125
 Db 185 DSK-FWCYIFKAKKSPD-----FC-----STPACTKEKECYTGKGLD 222

RESULT 4

08B099 PRELIMINARY; PRT; 516 AA.
 AC 08B099.
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Similar to plasminogen activator, tissue.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=SKIN;
 RA Strauberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.

CC EMBL; BC002795; AAH02795.1; -.
 DR HSSP; P00750; 1ASH.
 DR MEROPS; S01.232; -.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00008; EGF_1.
 DR Pfam; PF00051; kringle; 2.
 DR PRINTS; PR00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR ProDom; PD000395; Kringle; 2.
 DR SMART; SM00181; EGF_1.
 DR SMART; SM00001; EGF_1.
 DR SMART; SM00130; KR; 2.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_2.
 DR PROSITE; PS50070; KRINGLE_2; 2.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.


```

DE Placminogen precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Browne M.J., Chapman C.G., Dodd I., Carey J.E., Lawrence G.M.P.,
RT "Expression of recombinant human plasminogen and aglycoplasminogen in
RT HeLa cells." 0:0-0(1991).
RL Fldrnolysis 0:0-0(1991).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPsin FAMILY.
DR EMBL; M74220; AAA36451.1; -.
DR HSRP; P00747; 2PK4.
DR MEROPS; S01.233; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; Kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 5.
DR SMART; SM00473; PAN AP; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 5.
DR PROSITE; PS00070; KRINGLE_2; 5.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
DR Hydrolase; Serine protease; Signal.
FT SIGNAL 1 19
FT CHAIN 20 810 PLASMINOGEN.
SQ SEQUENCE 810 AA; 90555 MW; B05C7D4BD020B3C CRC64;

Query Match 10.2%; Score 142; DB 4; Length 810;
Best Local Similarity 37.2%; Pred. No. 0.00035;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

QY 24 CFMDNGHLYREDQTSPPAGRLCLNWLDA-----QGLASAPVSGAGNHSYCRNPDEDPG 78
DB 103 CKTGNGKRYRGSMKTKYNGKITCKMSTSPHRPRFSPTATPSEGL-EENYCRNPDDNDPQG 161
QY 79 PWCYVSGAVPEKR--PCEDLRCE 102
DB 162 PWCYTTD---PEKRYDCILCEE 163

RESULT 8
Q25101 PRELIMINARY; PRT; 385 AA.
ID Q25101;
AC Q25101;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Serine proteinase.
GN HMRSP;
OS Herdmania momus.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Scolidobranchia; Eurythoe; Herdmania.
OX NCBI_TaxID=7733;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=CURVATA;
RA Arnold J.M., Kennett C., Lavlin M.F.;
RT "Transient expression of a novel serine protease in the ectoderm of

```

```

RT the ascidian Herdmania momus during development.";
RL Dev. Genes Evol. 206:455-465(1997).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPsin FAMILY.
DR EMBL; U6317; AAB6650.1; -.
DR HSRP; P00763; IDPO.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
DR Hydrolase; Serine protease.
SQ SEQUENCE 385 AA; 42935 MW; BFB1D0505232BEA0 CRC64;

Query Match 10.1%; Score 139.5; DB 5; Length 385;
Best Local Similarity 20.2%; Pred. No. 0.00023;
Matches 70; Conservative 47; Mismatches 127; Indels 103; Gaps 14;

QY 1 LLAW--VOAFVSNMLAEAYSGGCF--WDNGHLYREDQTSPPAGRLCLNWLDAOS---- 53
DB 7 LVIMTLINGFVSN-----SECFDENPESYGALISITLGGFCGSM-DLQIFPHG 56
QY 54 --GLASAPVSGAGNHSYCRNPDEDPGWCYSGE-----AGVEKRPCEDLR 99
DB 57 KYTSNPNYNSGLAGNNYCRNPDDMGWGWGVCYTNEFMKWDYCDIPICSNPVTLPSTIE 116
QY 100 CPETT---SQALPAFTTEIQEASGEG--GADEVQ----- 128
DB 111 CGKTEPLSDATKGYDKSKAKTNPLHVGDTVTGHSIPWVSRLKRLRHPCGSI 176
QY 129 -----VFAPNALPASEAQAQPVGISQVRNMSKEDGLTIGVL----- 172
DB 177 LNRWILTAHCIRKPOQPKYLAIGDYDRIQYDFSEMK---VGRILFNHEKNPT 232
QY 173 --GITMVTIILAGAILIGSYKRGDLKEQHDQVC-----EREMQRITL 217
DB 233 FENDITLKMWDTSISATIFGQSVPPANKVPAAKSKIVSGMDTKGTQDVKLQVTL 292
QY 218 PLSAFTNPTCEIVDEKTV---VHTSQTPVDPOGSGTPLMGQAGTP 260
DB 229 PWSF--KLCKKLYSKVGAAPVFTSLCAAVKKGKDGSCGSGSGGP 317

RESULT 9
P70521 PRELIMINARY; PRT; 716 AA.
ID P70521;
AC P70521;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Macrophage stimulating protein precursor.
GN MSP;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RC MEDLINE=97011126; PubMed=8658136;
RA Ohsishi K., Iwama A., Matsuno K., Ezaki T., Sakamoto O., Hamaguchi I.,
RA Takasu N., Suda T.;
RT "Molecular cloning of Rat Macrophage-stimulating protein and its

```

RT involvement in the Male Reproductive System.";
RL Biochem. Biophys. Res. Commun. 227:273-280(1996).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
TRYPSIN FAMILY.
DR EMBL; X95096; CAA64473.1; -.
DR HSSP; P00747; IKRN.
DR MEROPS; S01.975; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringles.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; Kringles; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD00395; Kringles; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR PROSITE; PS00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR KMW; KMW0135; Serine protease; Signal.
FT SIGNAL 1 31
SQ SEQUENCE 716 AA; 80733 MW; 06B7DF3EF56D921F CRC64;
Query Match 9.9%; Score 137; DB 11; Length 716;
Best Local Similarity 27.3%; Pred. No. 0.0084;
Matches 48; Conservative 12; Mismatches 56; Indels 60; Gaps 8;
QY 1 LLAWVQAFVSNMLALAEVAGS-----GCFMDNGHLYREDQGTSPAPGLACLMW 48
DB 80 LLAWVQ-----HSLRAQLHHSCLPLFKQKYVTCIMDNGASRGTVATAGLPCQAM 134
QY 49 ---LDAQSLASAPVSGAGNSYCNEDPEDPGPCVYS----- 84
DB 135 SRFPNDHKYTPPKNGL-EENFCRNPDGDRPGMVCYTINSAVFQSGIKSCREAVCW 193
QY 85 -----GEAGVPEK-RPCE--DLRCPE-----TSQALPAFTTEIQEASEGP 122
DB 194 CNGEDYRGVEDYTESGREGCQKMDLHPHSHPEKPFDRALKDNYCRANDASERP 249

RESULT 10
Q90504 PRELIMINARY; PRT; 420 AA.
ID Q90504;
AC Q90504;
DT 01-NOV-1996 (TREMblrel. 01. Created)
DT 01-NOV-1996 (TREMblrel. 01. Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20. Last annotation update)
DE Thrombin.
OS Eptatretus stoutii (Pacific hagfish).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Myxiniiformes;
OC Myxiniidae; Eptatretinae; Eptatretus.
OX NCBI_TaxID=7765;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=92212913; PubMed=155783;
RA Banfield D.K., Macgillivray R.T.;
RT "Partial characterization of vertebrate prothrombin cDNAs:
RT amplification and sequence analysis of the B chain of thrombin from
RT nine different species";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=94223694; PubMed=7513365;
RA Banfield D.K., Irwin D.M., Walz D.A., Macgillivray R.T.;
RT "Evolution of prothrombin: isolation and characterization of the cDNAs
RT encoding chicken and hagfish prothrombin.";

RL J. Mol. Evol. 38:177-187(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Banfield D.K.;
RL Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
TRYPSIN FAMILY.
DR EMBL; M81393; AAA21620.1; -.
DR HSSP; P00734; IUVS.
DR MEROPS; S01.217; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringles.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; Kringles; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD00395; Kringles; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR KMW; KMW0135; Serine protease.
FT SIGNAL 1 420
SQ SEQUENCE 420 AA; 47888 MW; 64522AA21A57867A CRC64;
Query Match 9.8%; Score 136.5; DB 13; Length 420;
Best Local Similarity 26.5%; Pred. No. 0.0047;
Matches 39; Conservative 16; Mismatches 77; Indels 15; Gaps 4;
QY 24 CFMDNGHLYREDQGTSPAPGLACLMWLAQSLASAPVSGAG-NHSYCRNPDGPPGMCY 82
DB 17 CYRRSGRDYRGDNLITWTKGKPLWRGYSNPLPSQPTTNGLTSTNCRNPDGDSGCWCY 76
QY 83 VSGEAGVP---EKRPCEDLRCPEPTTSQALPAFTTEIQEASRGPADEVQAPANALPA 138
DB 77 TKGVEGTVDVYQCLNYCE-----SGDIFEGVTDEVOISGRSGAAEKTLFPNPKTRGN 129
QY 139 RSEAAAVOPVIGISQVRMNSKEKKDL 165
DB 130 GBECCGKPMPELQK--NDRSEDEL 153

RESULT 11
Q90503 PRELIMINARY; PRT; 812 AA.
ID Q90503;
AC Q90503;
DT 01-MAY-2000 (TREMblrel. 13. Created)
DT 01-MAY-2000 (TREMblrel. 13. Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20. Last annotation update)
DE Plasminogen protein precursor (EC 3.4.21.7).
GN PLASMINOGEN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=91250378; PubMed=1645711;
RA Kanalas J.J., Makker S.P.;
RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a
RT receptor site for plasminogen";
RL J. Biol. Chem. 266:10825-10829(1991).

```

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: AJ242649; CAB64014.1; -.
DR HSSP: P00747; 1PMK.
DR MEROPS: S01.233; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan_app.
DR InterPro: IPR001254; Ser_protease_Try.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00051; Kringle; 5.
DR Pfam: PF00024; PAN; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle; 5.
DR SMART: SM00130; KR; 4.
DR SMART: SM00473; PAN; 1.
DR SMART: SM00020; Tryp_Spc; 1.
DR PROSITE: PS00021; KRINGLE_1; 5.
DR PROSITE: PS50070; KRINGLE_2; 5.
DR PROSITE: PS00338; SOMATOTROPIN_2; UNKNOWN_1.
DR PROSITE: PS50240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
DR KMW: KMW00135; Serine protease; Signal.
FT SIGNAL 1 19
FT CHAIN 20 812 PLASMINOGEN.
SQ SEQUENCE 812 AA; 90535 MW; 8C703C51410EBC9E CRC64;

Query Match 9.7%; Score 135; DB 11; Length 812;
Best Local Similarity 30.9%; Pred. No. 0.0015;
Matches 34; Conservative 14; Mismatches 48; Indels 14; Gaps 5;

QY 24 CFWDNGHLYREDQTSPPAGRLCLNWLDA-----QSGLASAPVSGAGNHSYCRNPDEPRG 78
DB 376 CYQGNGKSRGTSSTTGKCKQSGVSMTPSHSKTPANPPDGL-ENMYCENPDNDRG 434
QY 79 PNCYVSGEAGVPERK--PCEDLRCPETTSQALPAFTTEIGASEGSGADE 126
DB 435 PNCFTTD---PSVWMEYCNKRCSEFTGGV--AESAIIVQVSAAGTSE 478

RESULT 12
Q9BZM1 PRELIMINARY; PRT; 395 AA.
ID Q9BZM1;
AC Q9BZM1;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Neocatal thrombolytic agent alpha-form (Fragment).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Dou D.;
RT "A brain-type plasminogen activator.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: AF260825; AAK11956.1; -.
DR HSSP: P00750; 1PK2.
DR MEROPS: S01.232; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000083; Fibrinctnl.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00039; tnl; 1.
DR Pfam: PF00051; Kringle; 1.
DR Pfam: PF00089; trypsin; 1.

```

```

DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle; 1.
DR SMART: SM00058; FN1; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; Tryp_Spc; 1.
DR PROSITE: PS01253; FIBRONECTIN_1; 1.
DR PROSITE: PS00021; KRINGLE_1; UNKNOWN_1.
DR PROSITE: PS50070; KRINGLE_2; 1.
DR PROSITE: PS50240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
DR KMW: KMW00135; Serine protease.
FT NON_TER 395
FT SEQUENCE 395 AA; 44323 MW; 3FBD4A2F0B7C11C8 CRC64;

Query Match 9.7%; Score 134.5; DB 4; Length 395;
Best Local Similarity 32.3%; Pred. No. 0.00066;
Matches 32; Conservative 13; Mismatches 33; Indels 21; Gaps 4;

QY 3 AMVOAFIVSNML-----LAEAVSGGCFWDNGHLYREDQTSPPAGRLCLNWL-----48
DB 18 SWLPVLRNRYEYCMNCGRAQCSGNSDCYFGNGSAYRGTHLTJESGASCLPMNSMIL 77
QY 49 ---LDQSGLASAPVSGAGNHSYCRNPDEPRGWCYV 83
DB 78 IGKYYTAQN--PSAQLGKGNKYNCRNPDGAK-PWCHV 113

RESULT 13
Q90865 PRELIMINARY; PRT; 704 AA.
ID Q90865;
AC Q90865;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Hepatocyte growth factor-like/macrophage stimulating protein.
GN HGFI/MSP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_Taxid=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver.
RX MEDLINE=96029010; PubMed=7554499;
RA Thery C., Sharpe M.J., Bailey S.J., Stern C.D., Cheraidi E.;
RT "Expression of HGF/SF, HGFI/MSP and c-met suggests new functions
RT during early chick development.";
RT Dev. Genet. 17:90-101(1995).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: X84043; CAAS9862.1; -.
DR HSSP: P00747; 1CEA.
DR MEROPS: S01.977; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan_app.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00051; Kringle; 4.
DR Pfam: PF00024; PAN; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle; 4.
DR SMART: SM00130; KR; 4.
DR SMART: SM00473; PAN; 1.
DR SMART: SM00020; Tryp_Spc; 1.
DR PROSITE: PS00021; KRINGLE_1; 4.
DR PROSITE: PS50070; KRINGLE_2; 4.
DR PROSITE: PS50240; TRYPSIN_DOM; 1.

```

KM Hydrolyase; Serine protease.
SQ SEQUENCE 704 AA; 79341 MW; CAB0D8C41367C37 CRC64;
Query Match 9.7%; Score 134.5; DB 13; Length 704;
Best Local Similarity 28.8%; Pred. No. 0.0014;
Matches 32; Conservative 11; Mismatches 45; Indels 23; Gaps 4;
QY 24 CFWMNGHLYREDQTSPPAPGLRCINWLDASGLASAP-----VSGAGNHSYCRNPED 75
DB 108 CIYANGTSYGRDTPRGLRQHW-----QATPHDHRFLPSLRNGLSENYCRNPED 161
QY 76 PRGWCYV-----SGAGVPEKRPCEDLRCPEITTSQALPATTEIOBASE 120
DB 162 KRGWCYTVDPNVVHOSGCI--KCEDAVCMTCNGEDYRGVFDHTESGTE 209
RESULT 14
Q91XG8 PRELIMINARY; PRT; 716 AA.
AC Q91XG8;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE Hepatocyte growth factor-like.
GN HGFL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Strauberg R.;
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; BC010551; AA010551.1; -
DR MGI; MGI:96080; HGFL.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR01254; Ser.protease_Try.
DR Pfam; PF00051; Kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR Pfam; PF00035; Kringle; 4.
DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
KW Hydrolyase; Serine protease.
SQ SEQUENCE 716 AA; 80693 MW; 12474C48A7D4B46D CRC64;
Query Match 9.7%; Score 134; DB 11; Length 716;
Best Local Similarity 28.2%; Pred. No. 0.0016;
Matches 42; Conservative 9; Mismatches 46; Indels 52; Gaps 7;
QY 1 LLAWQAFLVSNMLAEAYSG-----GCFWNGHLYREDQTSPPAPGLRCINW 48
DB 80 LLPTWQ-----HSHTQLYHSLCHLFQKKDYRTICIMONGVSYRGTVARTAGLPCQAW 134
QY 49 ---LDAOSGLASAVSAGAHNSYCRNPDEDPRGWCYV----- 84
DB 135 SRPFNDHKYTPPKNGI-ENFCRNPDGPRGWCYTTNRSVRQSGIKTCREAVCVL 193
QY 85 -----GEAGVPEK-RPCE--DLRCPET 103
DB 194 CNGEDYRGEVDVTESGRECRWDLQHPHS 222
RESULT 15
Q9PU78 PRELIMINARY; PRT; 313 AA.
AC Q9PU78;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)

DE Hepatocyte growth factor-like protein (Fragment).
OS Crocodylus niloticus (Nile crocodile) (African crocodile).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylidae; Crocodylinae; Crocodylus.
OX NCBI_TaxID=6501;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=20022983; PubMed=1055283;
RA Hughes S., Zelus D., Mouchiroud D.;
RT Warm-blooded isochore structure in Nile crocodile and turtle.";
RL Mol. Biol. Evol. 16:1521-1527(1999).
DR EMBL; AJ011396; CAB56422.1; -
DR HSSP; P00747; IHPV.
DR MEROPS; S01.977; -
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR01254; Ser.protease_Try.
DR Pfam; PF00051; Kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00019; KRINGLE.
DR ProDom; PD00035; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRY_SP; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
KW Hydrolyase; Serine protease.
FT NON_TER 1 1
FT NON_TER 313 313
SQ SEQUENCE 313 AA; 34793 MW; 8E084704958B5AA2 CRC64;
Query Match 9.5%; Score 131.5; DB 13; Length 313;
Best Local Similarity 29.8%; Pred. No. 0.00092;
Matches 28; Conservative 11; Mismatches 42; Indels 13; Gaps 4;
QY 24 CFWMNGHLYREDQTSPPAPGLRCINWLDASGLAS-APVSGAGNH--SYCRNPDDPRGP 79
DB 30 CYHNGELYRGHTSKTRKGVTCQKMSQSPHVPQISPTTHPAHLIDENYCRNPDDSHGP 89
QY 80 WCYVSGE-----AGVPEKRPCEDLRCPEITTSQA 107
DB 90 WCYIMDPRTPPDYCI--KPCAGDKIPSVLENA 120

Search completed: April 7, 2003, 08:43:28
Job time : 35 secs

4

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: April 7, 2003, 06:34:02 ; Search time 37 Seconds

943,558 Million cell updates/sec

Title: US-10-057-951-2_COPY_2_263

Sequence score: 1388
1 LIAWVQAFIVSNMLAEAYG.....PVDPGSTPTLMGQAGTTPGA 262

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

1: /SID82/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID82/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID82/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID82/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID82/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID82/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID82/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID82/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID82/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID82/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID82/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID82/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID82/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID82/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID82/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID82/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID82/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID82/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID82/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID82/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
21: /SID82/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
22: /SID82/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
23: /SID82/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1388	100.0	263	20	AAV05219
2	1388	100.0	263	22	AAH05769
3	1388	100.0	263	22	AAH00100
4	1385	99.8	263	23	AAH06149
5	1379	99.4	263	21	AAH04317
6	1374	99.0	263	22	AAH093748
7	1162.5	83.8	286	20	AAI05520
8	322	23.2	66	22	ABH37905
9	322	23.2	66	22	ABH23159
10	322	23.2	66	22	AAH58537

11	322	23.2	66	22	AAH71037
12	322	23.2	66	22	AAH18800
13	322	23.2	66	22	AAH31314
14	322	23.2	66	23	ABG40828
15	301	21.7	56	20	AAH12615
16	300	21.6	55	20	AAH12397
17	196	14.1	39	19	AAH72641
18	192	13.8	39	19	AAH72640
19	160.5	11.6	527	13	AAH20217
20	160.5	11.6	527	13	AAH20220
21	160.5	11.6	527	13	AAH20218
22	160.5	11.6	527	13	AAH20219
23	160.5	11.6	527	13	AAH20219
24	160.5	11.6	527	13	AAH20219
25	160.5	11.6	527	13	AAH20219
26	159.5	11.5	472	10	AAH23407
27	159.5	11.5	527	13	AAH24416
28	159.5	11.5	527	13	AAH20222
29	159.5	11.5	527	13	AAH20221
30	157.5	11.3	527	19	AAH54154
31	157.5	11.3	527	19	AAH54157
32	157.5	11.3	562	10	AAH94379
33	157.5	11.3	562	10	AAH94380
34	156.5	11.3	439	16	AAH68851
35	156.5	11.3	483	16	AAH70889
36	156.5	11.3	483	16	AAH70877
37	156.5	11.3	483	16	AAH70878
38	156.5	11.3	483	16	AAH70879
39	156.5	11.3	483	16	AAH70880
40	156.5	11.3	483	16	AAH70881
41	156.5	11.3	483	16	AAH70882
42	156.5	11.3	483	16	AAH70883
43	156.5	11.3	483	16	AAH70884
44	156.5	11.3	483	16	AAH70885
45	156.5	11.3	483	16	AAH70886

ALIGNMENTS

RESULT 1	
AAV05219	
ID	AAV05219 standard; Protein, 263 AA.
XX	
AC	AAV05219;
XX	
DT	17-JUN-1999 (first entry)
XX	
DE	Kringle1 protein sequence.
XX	
KM	Kringle1; cancer; inflammation; autoimmunity; rheumatoid arthritis; AIDS;
KM	CNS inflammation; cerebellar degeneration; Alzheimer's disease; asthma;
KM	Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
KM	neurological abnormality; ischaemia reperfusion injury; ischemic injury;
KM	cardiovascular disease; kidney disease; liver disease; aplastic anemia;
KM	myocardial infarction; hypotension; hypertension; allergy; infection;
KM	myelodysplastic syndrome; haematologic abnormality; diagnosis; therapy;
KM	male pattern baldness.
XX	
OS	Homo sapiens.
XX	
PN	WO931786-A1.
XX	
PD	11-MAR-1999.
XX	
PF	02-SEP-1998; 98MO-US18270.
XX	
PR	01-SEP-1998; 98US-0144889.
XX	
PR	02-SEP-1997; 97US-0056037.
XX	
PA	(SMK) SMITHKLINE BEECHAM CORP.
XX	
PI	Albone EF, Kikly KK;

```

XX WPI: 1999-214707/18.
DR N-PSDB; AAK28354.
XX
PT New kringle1 polypeptides and polynucleotides
XX
PS Claim 1; Page 31-32; 42pp; English.
XX
CC This sequence is a Kringle1 polypeptide of the invention.
CC The Kringle1 polypeptides (1) are used to screen for agonists and
CC antagonists. Agonists are used to treat subjects in need of enhanced
CC activity or expression of (1). Antagonists are used to treat subjects
CC having need to inhibit the activity or expression of (1). The methods can
CC be used to treat conditions such as cancer, inflammation, autoimmunity,
CC allergy, asthma, rheumatoid arthritis, CNS inflammation, cerebellar
CC degeneration, Alzheimer's disease, Parkinson's disease, multiple
CC sclerosis, amyotrophic lateral sclerosis, head injury damage and other
CC neurological abnormalities, ischaemia reperfusion injury, cardiovascular
CC infarction, hypotension, hypertension, AIDS, myelodysplastic syndromes
CC and other haematologic abnormalities, aplastic anaemia, male pattern
CC baldness, and bacterial, fungal, protozoan and viral infections. The
CC Kringle1 polypeptides may also be used to generate antibodies.
CC Determining the presence or absence of mutations in, and analysing for
CC the presence or absence of expression of, Kringle1 polynucleotides can be
CC used to diagnose a disease or susceptibility to a disease related to
CC expression or activity of Kringle1 proteins. The polynucleotides may also
CC be used for chromosome identification, and mapping.
XX
SQ Sequence 263 AA;
XX
Query Match 100.0%; Score 1388; DB 20; Length 263;
Best Local Similarity 100.0%; Pred. No. 2, 1e-13;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 LLAATGATLVSNNMLLAAYVSGGCTWDNGHLYREDQTSPPAGRCINMLDQSLASAPV 60
DB 2 LLAATGATLVSNNMLLAAYVSGGCTWDNGHLYREDQTSPPAGRCINMLDQSLASAPV 61
QY 61 SGAGHASTYCRNPDDPRGWCYSGAGVPEKPCEDLRCPETTSQALPAFTTEIQASR 120
DB 62 SGAGHASTYCRNPDDPRGWCYSGAGVPEKPCEDLRCPETTSQALPAFTTEIQASR 121
QY 121 GFGADGVQVFPANALPARSEAAVQPIGISOVRNNSKEKKDGLGYVLGITMMV11 180
DB 122 GFGADGVQVFPANALPARSEAAVQPIGISOVRNNSKEKKDGLGYVLGITMMV11 181
QY 181 IAIAGIILGYSYKRGKDLKEQDKVCEMRORITPLSAFTNPCEIVDEKTVVHTS 240
DB 182 IAIAGIILGYSYKRGKDLKEQDKVCEMRORITPLSAFTNPCEIVDEKTVVHTS 241
QY 241 QTPVDPQEGSTPLMGQAGTFGA 262
DB 242 QTPVDPQEGSTPLMGQAGTFGA 263
XX
RESULT 2
AAW87769
ID AAW87769 standard; Protein; 263 AA.
XX
AC AAW87769;
XX
DT 29-MAR-1999 (first entry)
XX
DE Human tissue plasminogen activator-like protease t-PALP.
XX
KM Tissue plasminogen activator-like protease; t-PALP; human;
KM circulatory system-related disorder; blood clotting; stroke;
KM thrombosis; peripheral arterial occlusion; pulmonary embolism;
KM myocardiobrombosis; diagnosis; therapy.
XX
OS Homo sapiens.
XX

```

```

FH Key Location/Qualifiers
FH Peptide 1..21
FH Protein 22..263
FH Domain 25..84
FH Domain 85..263
FH Domain 85..263
FH Peptide 22..31
FH Peptide 35..44
FH Peptide 71..81
FH Peptide 91..107
FH Peptide 119..128
FH Peptide 138..147
FH Peptide 155..167
FH Peptide 193..203
FH Peptide 206..215
FH Peptide 227..237
FH Peptide 243..252
FH Peptide /note="epitope-bearing region"
FH W09854199-A1.
FH 03-DEC-1998.
FH 27-MAY-1998; 98W0-US10728.
FH 28-MAY-1997; 97US-0048000.
FH (HUMA-) HUMAN GENOME SCI INC.
FH Edner R, Moore PA, Ruben SM;
FH WPI: 1999-070207/06.
FH N-PSDB; AAV99636.
FH
PS Claim 1; Page 56-57; 76pp; English.
XX
CC This is the amino acid sequence of tissue plasminogen activator-like
CC protease (t-PALP), a novel member of the serine protease family
CC that shares sequence homology to human tissue plasminogen activator
CC (see AAW87769). The t-PALP sequence was deduced from a cDNA clone
CC (see AAV99636) derived from activated monocytes. The 2.5 kb t-PALP
CC message has also been detected in heart, brain, lung, placenta,
CC liver, skeletal muscle, kidney, pancreas, spleen, thymus, prostate,
CC testis, ovary, small intestine, colon and peripheral blood
CC leukocytes. Isolated nucleic acids encoding amino acids -21 to
CC 242, -20 to 242, 1-242, 4-63 (kringle domain) and 64-242 (protease
CC domain) of t-PALP, or encoding epitope-bearing portions of t-PALP,
CC are also claimed, as are recombinant vectors, host cells, and
CC methods for producing t-PALP polypeptides. t-PALP may be used to
CC detect and treat disorders related to the circulatory system, and
CC to identify agonists and antagonists of t-PALP activity. The
CC homology between t-PALP and tPA indicates that t-PALP may be
CC involved in the regulation of normal and abnormal clotting
CC in e.g. stroke, deep-vein thrombosis, peripheral arterial
CC occlusion, pulmonary embolism and myocardiobrombosis.
XX
SQ Sequence 263 AA;

```

Query Match	100.0%	Score 1388	DB 20	Length 263
Best Local Similarity	100.0%	Pred. No. 2,1e-119		
Matches 262	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	LLAAVQAATLVSMMLAAEYVSGSGCFCWMDNCHLYREDQTSAPAGLRCLNMLDAOSGLASAPV	60	
Db	2	LLAAVQAATLVSMMLAAEYVSGSGCFCWMDNCHLYREDQTSAPAGLRCLNMLDAOSGLASAPV	61	
QY	61	SGAGNHSYCRNDEDEPRGWCYVSGEAGVPEKRPCEDLACPETTSQALPAFTTEIOEASE	120	
Db	62	SGAGNHSYCRNDEDEPRGWCYVSGEAGVPEKRPCEDLACPETTSQALPAFTTEIOEASE	121	
QY	121	GGGADEVQYFAFANALLPARSEAAAVQPVIGISQRYVMSKEKKDGGTIGYVIGITMVTII	180	
Db	122	GGGADEVQYFAFANALLPARSEAAAVQPVIGISQRYVMSKEKKDGGTIGYVIGITMVTII	181	
QY	181	IAIGAGIILIGSYKRGKDLKEQHDQVCEREMQRTILPLSAFLNPICELVDEKTVVHTS	240	
Db	182	IAIGAGIILIGSYKRGKDLKEQHDQVCEREMQRTILPLSAFLNPICELVDEKTVVHTS	241	
QY	241	QTPVDPQEGSTPLMGQAGTPGA	262	
Db	242	QTPVDPQEGSTPLMGQAGTPGA	263	

RESULT 3	
AAE00300	
ID	AAE00300 standard; Protein, 263 AA.
XX	
AC	AAE00300;
XX	
DT	13-JUN-2001 (first entry)
XX	
DE	Human tissue-plasminogen activator-like protease (t-PALP).
XX	
KW	Human; tissue-plasminogen activator-like protease; t-PALP;
KW	therapy; vascular disease; stroke; deep vein thrombosis; keloid; asthma;
KW	arterial occlusion; blood coagulation disorder; cerebroprotective;
KW	autoimmune system disorder; human immunodeficiency syndrome; cytostatic;
KW	rheumatoid arthritis; graft-versus-host disease; thyroiditis; candida;
KW	insulin dependent diabetes; inflammatory eye disease; allergy; antiviral;
KW	cardiovascular disease; heart disease; arrhythmia; myocardial ischaemia;
KW	hyperproliferative disorder; hypertrophic scar; neurological disease;
KW	Creutzfeldt-Jakob syndrome; neurodegenerative disorder; antiinflammatory;
KW	Alzheimer's disease; Parkinson's disease; immunosuppressive; antifungal;
KW	infectious disease; drug screening; gene therapy; neuroprotective;
KW	cancer; ophthalmological; antibacterial; vulnerability.
XX	
OS	Homo sapiens.
XX	
FH	
Key	Location/Qualifiers
FT	1..165
FT	/note= "binds to FLAG polypeptide to form
FT	t-PALP-FLAG fusion protein"
FT	
FT	1..21
FT	/label= Signal_peptide
FT	
FT	4..63
FT	/label= Kringle_domain
FT	
FT	12..21
FT	/note= "Conserved region"
FT	
FT	22..263
FT	/note= "Human mature tissue-plasminogen activator-like
FT	protease (t-PALP); Binds to FLAG polypeptide to form
FT	t-PALP-FLAG fusion protein"
FT	
FT	22..38
FT	/note= "Conserved region"
FT	
FT	22..31
FT	/note= "Epitope-bearing portion"
FT	
FT	35..44
FT	/note= "Epitope-bearing portion"
FT	
FT	39..49
FT	/note= "Conserved region"

FT	Region	50..62	/note= "Conserved region"
FT	Region	63..84	/note= "Conserved region"
FT	Domain	64..242	/label= Protease_domain
FT	Region	71..81	/note= "Epitope-bearing portion"
FT	Region	85..97	/note= "Conserved region"
FT	Region	91..107	/note= "Epitope-bearing portion"
FT	Region	100..118	/note= "Conserved region"
FT	Region	119..128	/note= "Epitope-bearing portion"
FT	Region	119..127	/note= "Conserved region"
FT	Region	128..143	/note= "Conserved region"
FT	Region	138..147	/note= "Epitope-bearing portion"
FT	Region	146..163	/note= "Conserved region"
FT	Region	155..167	/note= "Epitope-bearing portion"
FT	Region	164..180	/note= "Conserved region"
FT	Region	186..200	/note= "Conserved region"
FT	Region	193..203	/note= "Epitope-bearing portion"
FT	Region	201..220	/note= "Conserved region"
FT	Region	206..215	/note= "Epitope-bearing portion"
FT	Region	221..236	/note= "Conserved region"
FT	Region	227..237	/note= "Epitope-bearing portion"
FT	Region	237..248	/note= "Conserved region"
FT	Region	243..252	/note= "Epitope-bearing portion"
FT	Region	249..263	/note= "Conserved region"
PN	MO200125252-A1.		
PD	12-APR-2001.		
PF	03-OCT-2000; 2000MO-US27239.		
PR	04-OCT-1999; 99US-0411977.		
PA	(HUMA-) HUMAN GENOME SCT INC.		
PI	Moore PA, Ruben SM, Ebner R;		
DR	WPI; 2001-235402/24.		
DR	N-PFSD; AAD03460.		
XX	New (gene encoding and antibody immunospecific for a)		
XX	tissue-Plasminogen activator-like protease, useful for the diagnosis		
XX	and treatment of (cardio)vascular diseases, hyperproliferative		
XX	disorders, immune system disorders and cancers -		
PS	Claim 17; Fig 1; 323pp; English.		
XX	The present amino acid sequence is HMSIB42 clone human		
XX	tissue-Plasminogen activator-like protease (t-PALP). The t-PALP		
XX	sequence and their (anti)agonists are useful for the diagnosis and		
XX	treatment of vascular diseases e.g. stroke, deep vein thrombosis and		
CC			

Claim 17; Fig 1; 323pp; English.

The present amino acid sequence is HMS142 clone human tissue-plasminogen activator-like protease (t-PALP). The t-PALP sequence and their (ant)agonists are useful for the diagnosis and treatment of vascular diseases e.g. stroke, deep vein thrombosis and

arterial occlusion, blood coagulation disorders, (auto)immune system disorders e.g. human immunodeficiency syndrome, rheumatoid arthritis, graft-versus-host disease, thyroiditis, insulin dependent diabetes and inflammatory eye disease, allergic reactions e.g. asthma, cardiovascular diseases e.g. heart disease, arrhythmia and myocardial ischaemia, hyperproliferative disorders, cancers, hypertrophic scars and keloids, neurological diseases e.g. Creutzfeldt-Jakob syndrome, neurodegenerative disorders e.g. Alzheimer's disease and Parkinson's disease and infectious disease e.g. viral, bacterial and fungal infections. The t-PALP sequences are also useful for drug screening. The t-PALP nucleotides are useful as chromosome markers and are involved in gene therapy.

Sequence 263 AA;

Query Match 100.0%; Score 1388; DB 22; Length 263;

Best Local Similarity 100.0%; Pred. No. 2.1e-119;

Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLAWQAFVLSNMLAEAYSGGCFWNGHLYREDQTSAPAGRLCLMWLDQSGLASAPV 60

Db 2 LLAWQAFVLSNMLAEAYSGGCFWNGHLYREDQTSAPAGRLCLMWLDQSGLASAPV 61

QY 61 SGAGNHSYCRNPDEDPKPCVSGAGVPEKRPCEDLRCEITTSQALPAFTTEIQASE 120

Db 62 SGAGNHSYCRNPDEDPKPCVSGAGVPEKRPCEDLRCEITTSQALPAFTTEIQASE 121

QY 121 GPGADEVQVFPANALPARSEAAVOPVIGISQVRMNSKEKDLGTLGYVLGITMMVII 180

Db 122 GPGADEVQVFPANALPARSEAAVOPVIGISQVRMNSKEKDLGTLGYVLGITMMVII 181

QY 181 IAIAGGIIIGSYRKGDKLKEQHDQKVCEREMQRIITPLSAFTNPTEIYDEKTVVHTS 240

Db 182 IAIAGGIIIGSYRKGDKLKEQHDQKVCEREMQRIITPLSAFTNPTEIYDEKTVVHTS 241

QY 241 QTPVDPQEGSTPLMGQAGTPGA 262

Db 242 QTPVDPQEGSTPLMGQAGTPGA 263

RESULT 4

AAU86149 ID AAU86149 standard; Protein; 263 AA.

XX AAU86149;

DT 15-JUL-2002 (first entry)

XX Human PRO264 polypeptide.

DE Human; PRO; benign tumour; malignant tumour; lymphoid malignancy;

KW Leukemia; neuronal disorder; stromal disorder; blastocoealic disorder;

KW Inflammatory disorder; immune disorder; angiogenic disorder;

XX cytostatic; neuroprotective.

OS Homo sapiens.

XX WO200153486-A1.

PN 26-JUL-2001.

XX 11-FEB-2000; 2000WO-US03565.

PF 08-MAR-1999; 99WO-US05028.

PR 11-MAR-1999; 99US-123972P.

PR 11-MAY-1999; 99US-133459P.

PR 02-JUN-1999; 99WO-US12252.

PR 22-JUN-1999; 99US-140650P.

PR 22-JUN-1999; 99US-140653P.

PR 20-JUL-1999; 99US-144758P.

PR 26-JUL-1999; 99US-145698P.

PR 28-JUL-1999; 99US-146222P.

PR 17-AUG-1999; 99US-149395P.

PR 31-AUG-1999; 99US-151689P.

PR 01-SEP-1999; 99WO-US20111.

PR 15-SEP-1999; 99WO-US21090.

PR 30-NOV-1999; 99WO-US28313.

PR 01-DEC-1999; 99WO-US28301.

PR 01-DEC-1999; 99WO-US28314.

PR 05-JAN-2000; 2000WO-US08219.

XX (GENTH) GENENTECH INC.

PA Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KU;

XX Masters SA, Pan J, Pitti RM, Roy MA, Smith V, Stone DM;

PI Watanabe CK, Wood WI;

XX WPI; 2002-205567/26.

DR N-PSDB; ABR40275.

XX Thirty five nucleic acids encoding PRO polypeptides, useful for

PT treating benign or malignant tumours, leukemias and lymphoid

PT malignancies, inflammatory, angiogenic and immunologic disorders -

PS Claim 61; Fig 44; 302pp; English.

XX The present invention relates to the isolation of novel human PRO

CC polypeptides and the polynucleotide sequences encoding them. The

CC PRO polypeptides, agonists, antagonists or anti-PRO antibodies are

CC useful for treating benign or malignant tumours (e.g. renal, kidney,

CC bladder, breast, etc), leukemias and lymphoid malignancies, other

CC disorders such as neuronal, glial, astrocytal, hypothalamic, glandular,

CC macrophagal, stromal and blastocoealic disorders, inflammatory, immune

CC and angiogenic disorders. The polynucleotide sequences are also

CC useful in gene therapy. AU86128-AU86162 represent the human PRO

CC polypeptides of the invention.

XX Sequence 263 AA;

QY 241 QTPVDPQEGSTPLMGQAGTPGA 262

Db 242 QTPVDPQEGSTPLMGQAGTPGA 263

QY 1 LLAWQAFVLSNMLAEAYSGGCFWNGHLYREDQTSAPAGRLCLMWLDQSGLASAPV 60

Db 2 LLAWQAFVLSNMLAEAYSGGCFWNGHLYREDQTSAPAGRLCLMWLDQSGLASAPV 61

QY 61 SGAGNHSYCRNPDEDPKPCVSGAGVPEKRPCEDLRCEITTSQALPAFTTEIQASE 120

Db 62 SGAGNHSYCRNPDEDPKPCVSGAGVPEKRPCEDLRCEITTSQALPAFTTEIQASE 121

QY 121 GPGADEVQVFPANALPARSEAAVOPVIGISQVRMNSKEKDLGTLGYVLGITMMVII 180

Db 122 GPGADEVQVFPANALPARSEAAVOPVIGISQVRMNSKEKDLGTLGYVLGITMMVII 181

QY 181 IAIAGGIIIGSYRKGDKLKEQHDQKVCEREMQRIITPLSAFTNPTEIYDEKTVVHTS 240

Db 182 IAIAGGIIIGSYRKGDKLKEQHDQKVCEREMQRIITPLSAFTNPTEIYDEKTVVHTS 241

QY 241 QTPVDPQEGSTPLMGQAGTPGA 262

Db 242 QTPVDPQEGSTPLMGQAGTPGA 263

RESULT 5

AAB43237 ID AAB43237 standard; Protein; 263 AA.

XX AAB43237;

DT 08-FEB-2001 (first entry)

XX Human ORFX ORF3001 polypeptide sequence SEQ ID NO:6002.

DE Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;

KW vulnerable; antiproliferative; antiparkinsonian; nootropic; neuroprotective;

KW anticonvulsant; osteopathic; antiarrhythmic; immunosuppressant; cardiac;

KM immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KM hypotensive; dermatological; immunosuppressive; antiinflammatory;
KM antiviral; antibacterial; antifungal; antirheumatic; antihistoid;
KM antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KM neurodegenerative disorder; osteoarthritis; graft vs host disease;
KM cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KM cholesterol ester storage; systemic lupus erythematosus; infection;
KM severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KM allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KM bone damage; cartilage damage; antiinflammatory disease; coagulation;
KM thrombosis; contraceptive.
OS Homo sapiens.
XX W0200058473-A2.
XX
XX
XX 05-OCT-2000.
XX
XX 31-MAR-2000; 2000WO-US08621.
XX
XX 31-MAR-1999; 99US-0127607.
XX 02-APR-1999; 99US-0127636.
XX 05-APR-1999; 99US-0127728.
XX 30-MAR-2000; 2000US-0540763.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shimkets RA, Leach M;
XX
XX WPI; 2000-602362/57.
XX N-PSDB; AAC77446.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease -
XX
XX Claim 11; Page 5181-5182; 5507tp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vulnary;
XX antiproliferative; antiparisonian; nootropic; neuroprotective;
XX osteopathic; anticonvulsant; antiahrbic; immunosuppressant;
XX immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
XX antidiabetic; hypotensive; dermatological; immunosuppressive;
XX antiinflammatory; antibacterial; antiviral; antifungal; antineumatic;
XX antihyroid; and antianaemic. The sequences can be used for determining
XX the presence of or predisposition to, or preventing or treating
XX pathological conditions associated with an ORFX-associated disorder. The
XX nucleic acids can be used to express ORFX proteins in gene therapy
XX vectors. The proteins and nucleic acids may be used to treat cancers,
XX proliferative disorders, neurodegenerative disorders, osteoarthritis,
XX graft vs host disease, cardiovascular disease, diabetes mellitus,
XX hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
XX bacterial or fungal infection, malaria, autoimmune disorders, asthma,
XX allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
XX nocturnal haemoglobinuria, antiinflammatory disease; to enhance
XX coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ Sequence 263 AA;
Query Match 99.4%; Score 1379; DB 21; Length 263;
Best Local Similarity 99.2%; Pred. No. 1.4e-118;
Matches 260; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 121 GGADEVQYFAPANALPARSEAAVQPIGHSORVRNNSKEKDLGTLGYLGTWVLI 180
DB 122 GGADEVQYFAPANALPARSEAAVQPIGHSORVRNNSKEKDLGTLGYLGTWVLI 181
QY 181 IAIAGATIIIGSYKRGKDKIEQHDQKVCEREMORITPLSAFTNPTGCIYDEKTVVHTS 240
DB 182 IAIAGATIIIGSYKRGKDKIEQHDQKVCEREMORITPLSAFTNPTGCIYDEKTVVHTS 241
QY 241 QTPVDPQEGSTPLMGQAGNPGA 262
DB 242 QTPVDPQEGSTPLMGQAGNPGA 263
RESULT 6
AAM93748
ID AAM93748 standard; Protein; 263 AA.
XX
XX AAM93748;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human polypeptide, SEQ ID NO: 3727.
XX
XX Human; full length cDNA; cDNA synthesis; oligo-capping.
XX
XX Homo sapiens.
XX
XX EPI130094-A2.
XX
XX 05-SEP-2001.
XX
XX 07-JUL-2000; 2000EP-0114089.
XX
XX 08-JUL-1999; 99UP-0194486.
XX 11-JAN-2000; 2000JP-0118774.
XX 02-MAY-2000; 2000JP-0183765.
XX
XX (HELI-) HELIX RES INST.
XX
XX Oca T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,
XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI: 2001-524255/58.
XX N-PSDB; AAK94700.
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their
XX use in genetic manipulation -
XX
XX Claim 8; SEQ ID NO 3727; 1380bp + sequence listing; English.
XX
XX The invention relates to primers for synthesizing full length cDNA
XX clones. 830 cDNA molecules encoding a human protein have been
XX isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
XX molecules have been determined. Primers for synthesizing the full length
XX cDNA are useful for clarifying the function of the protein encoded by
XX the cDNA. The full length clones were obtained by construction of full
XX length enriched cDNA libraries that were synthesised by the oligo-capping
XX method. The primers enable the production of the full length cDNA easily
XX without any special methods. The present sequence is a polypeptide
XX encoded by a full length human cDNA of the invention.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ Sequence 263 AA;
Query Match 99.0%; Score 1374; DB 22; Length 263;
Best Local Similarity 99.2%; Pred. No. 4.1e-118;
Matches 260; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 61 SGAGNHSYCRNPDEDPGRGWCYSGAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASE 120
    |||
DB 62 SGAGNHSYCRNPDEDPGRGWCYSGAGVPEKRPCETTSQALPAFTTEIQEASE 121
QY 121 GPGADEVQVFAPANALPARSEAAAVQPIGTSQRVRNMSKEKDLGTLGYLGITMMVI 180
    |||
DB 122 GPGADEVQVFAPANALPARSEAAAVQPIGTSQRVRNMSKEKDLGTLGYLGITMMVI 181
QY 181 IAIAGGIILGYSYKRGKDLKEQHDQKVCEREMQRIITPLSAFTNPCEIIDEKTVVHTS 240
    |||
DB 182 IAIAGGIILGYSYKRGKDLKEQHDQKVCEREMQRIITPLSAFTNPCEIIDEKTVVHTS 241
QY 241 QTPVPDQEGSTPLMGQAGTPGA 262
    |||
DB 242 QTPVPDQEGSTPLMGQAGTPGA 263

RESULT 7
AA05220
ID AA05220 standard; Protein, 286 AA.
XX
AC AA05220;
XX
DT 17-JUN-1999 (first entry)
XX
DE Kringel protein sequence.
XX
KW Kringel; cancer; inflammation; autoimmunity; rheumatoid arthritis; AIDS;
KW CNS inflammation; cerebellar degeneration; Alzheimer's disease; asthma;
KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
KW neurological abnormality; ischaemia reperfusion injury; ischaemic injury;
KW cardiovascular disease; kidney disease; liver disease; aplastic anaemia;
KW myocardial infarction; hypertension; allergy; infection;
KW myelodysplastic syndrome; haematologic abnormality; diagnosis; therapy;
KW male pattern baldness.
XX
OS Homo sapiens.
XX
PN MO9911788-A1.
XX
PD 11-MAR-1999.
XX
PF 02-SEP-1998; 98WO-US18270.
XX
PR 01-SEP-1998; 98US-0144889.
XX
PR 02-SEP-1997; 97US-0056032.
XX
PA (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI Albion EF, Kikly KK;
XX
DR WPI; 1999-214707/18.
XX
DR N-PSDB; AAX28355.
XX
New kringel polypeptides and polynucleotides
XX
PS Claim 14; Page 33; 42pp; English.
XX
This sequence is a Kringel polypeptide of the invention.
CC The kringel polypeptides (I) are used to screen for agonists and
CC antagonists. Agonists are used to treat subjects in need of enhanced
CC activity or expression of (I). Antagonists are used to treat subjects
CC having need to inhibit the activity or expression of (I). The methods can
CC be used to treat conditions such as cancer, inflammation, autoimmunity,
CC allergy, asthma, rheumatoid arthritis, CNS inflammation, cerebellar
CC degeneration, Alzheimer's disease, Parkinson's disease, multiple
CC sclerosis, amyotrophic lateral sclerosis, head injury damage and other
CC neurological abnormalities, ischaemia reperfusion injury, cardiovascular
CC disease, kidney disease, liver disease, ischaemic injury, myocardial
CC infarction, hypertension, AIDS, myelodysplastic syndromes
CC and other haematologic abnormalities, aplastic anaemia, male pattern
CC baldness, and bacterial, fungal, protozoan and viral infections. The

```

```

CC kringel polypeptides may also be used to generate antibodies.
CC Determining the presence or absence of mutations in, and analysing for
CC the presence or absence of expression of, kringel polynucleotides can be
CC used to diagnose a disease or susceptibility to a disease related to
CC expression or activity of kringel proteins. The polynucleotides may also
CC be used for chromosome identification, and mapping.
SQ Sequence 286 AA;
Query Match 83.8%; Score 1162.5; DB 20; Length 286;
Best local similarity 87.3%; Pred. No. 1,3e-98;
Matches 226; Conservative 5; Mismatches 27; Indels 1; Gaps 1;
QY 1 LLAWQAFVLSNMLLAEAYSGGCFWMDNGHLYREDQTSAPGLRCIMWIDAGSGAPAV 60
    |||
DB 2 LLAWQAFVLSNMLLAEAYSGGCFWMDNGHLYREDQTSAPGLRCIMWIDAGSGAPAV 61
QY 61 SGAGNHSYCRNPDEDPGRGWCYSGAGVPEKRPCEDLRCPETTSQALPAFTTEIQEAS 119
    |||
DB 62 SGAGNHSYCRNPDEDPGRGWCYSGAGVPEKRPCEDLRCPETTSQALPAFTTEIQEAS 121
QY 120 EGPDADEVQVFAPANALPARSEAAAVQPIGTSQRVRNMSKEKDLGTLGYLGITMMVI 179
    |||
DB 122 EGPDADEVQVFAPANALPARSEAAAVQPIGTSQRVRNMSKEKDLGTLGYLGITMMVI 181
QY 180 IAIAGGIILGYSYKRGKDLKEQHDQKVCEREMQRIITPLSAFTNPCEIIDEKTVVHT 239
    |||
DB 182 IAIAGGIILGYSYKRGKDLKEQHDQKVCEREMQRIITPLSAFTNPCEIIDEKTVVHT 241
QY 240 SOTVPDQEGSTPLMGQAG 258
    |||
DB 242 SOTVPDQEGSTPLMGQAG 260

RESULT 8
ABB37905
ID ABB37905 standard; Peptide; 66 AA.
XX
AC ABB37905;
XX
DT 04-FEB-2002 (first entry)
XX
DE Peptide #5411 encoded by human foetal liver single exon probe.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
PN MO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00669.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human fetal liver -
XX
PS Claim 27; SEQ ID NO 30540; 639pp + sequence listing; English.

```

CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 66 AA;

Query Match 23.2%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 3.5e-22;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 TTSQLPAPFTTEIOEASGPGADEVQVFPANALPARSEAAAVQVIGISQVRMNSKEK 162
DB 1 TTSQLPAPFTTEIOEASGPGADEVQVFPANALPARSEAAAVQVIGISQVRMNSKEK 60
QY 163 KDLGTL 168
DB 61 KDLGTL 66

RESULT 9
AB23159
ID AB23159 standard; Protein; 66 AA.
XX
AC AB23159;
XX
DT 23-JAN-2002 (first entry)
XX
DE Protein #5158 encoded by probe for measuring heart cell gene expression.
XX
KW Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00666.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX
PS Claim 15; SEQ ID No 24929; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC AB23159-AB23159). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the

CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 66 AA;

Query Match 23.2%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 3.5e-22;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 TTSQLPAPFTTEIOEASGPGADEVQVFPANALPARSEAAAVQVIGISQVRMNSKEK 162
DB 1 TTSQLPAPFTTEIOEASGPGADEVQVFPANALPARSEAAAVQVIGISQVRMNSKEK 60
QY 163 KDLGTL 168
DB 61 KDLGTL 66

RESULT 10
AAMS5837
ID AAMS5837 standard; Protein; 66 AA.
XX
AC AAMS5837;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 30642.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX
PS Example 4; SEQ ID NO: 30642; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 66 AA;

Query Match 23.2%; Score 322; DB 22; Length 66;

Best Local Similarity 100.0%; Pred. No. 3.5e-22;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 TTSQLPATTETIOASBPGADDEVQVFAPANALPARSEAAAVQPVIGISQVRNMSKEX 162
DB 1 TTSQLPATTETIOASBPGADDEVQVFAPANALPARSEAAAVQPVIGISQVRNMSKEX 60

QY 163 KDLGTL 168
DB 61 KDLGTL 66

RESULT 11

AAAM71037
ID AAAM71037 standard; Protein; 66 AA.

AC AAAM71037;

DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed probe encoded protein SEQ ID NO: 31343.

KM Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukemia; lymphoma; myeloma.

OS Homo sapiens.

PN WO200157276-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00668.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX PA Penn SG, Hanzel DK, Chen W, Rank DR;

XX PI WPI; 2001-488900/53.

XX DR WPI; 2001-488900/53.

XX PT Human genome-derived single exon nucleic acid probes useful for

XX PT analyzing gene expression in human bone marrow -

XX PS Example 4; SEQ ID NO: 31343; 658bp + Sequence Listing; English.

XX CC The present invention provides a number of single exon nucleic acid

XX CC probes which are derived from genomic sequences expressed in bone marrow

XX CC bone marrow. They can be used to measure gene expression in bone marrow

XX CC samples, which may enable the improved diagnosis and treatment of cancers

XX CC such as lymphoma, leukemia and myeloma. The present sequence is a

XX CC protein encoded by one of the probes of the invention.

XX CC

XX CC

XX CC

RESULT 12

AAAM18800
ID AAAM18800 standard; Protein; 66 AA.

AC AAAM18800;

DT 12-OCT-2001 (first entry)

DE Peptide #5234 encoded by probe for measuring cervical gene expression.

KM Probe; human; microarray; gene expression; cervical epithelial cell;

XX KW cervical cancer.

XX OS Homo sapiens.

PN WO200157278-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00670.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX PA Penn SG, Hanzel DK, Chen W, Rank DR;

XX PI WPI; 2001-488901/53.

XX DR WPI; 2001-488901/53.

XX PT Human genome-derived single exon nucleic acid probes useful for

XX PT analyzing gene expression in human cervical epithelial cells -

XX PS Claim 27; SEQ ID NO 23626; 487bp; English.

XX CC The present invention relates to human single exon nucleic acid probes

XX CC (SENP: see AA10068-AA128459). The present sequence is a peptide encoded

XX CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs

XX CC can be used to produce a single exon microarray, which can be used for

XX CC measuring human gene expression in a sample derived from human cervical

XX CC epithelial cells. By measuring gene expression, the probes are therefore

XX CC useful in grading and/or staging of diseases of the cervix, notably

XX CC cervical cancer.

XX CC Note: The sequence data for this patent did not form part of the printed

XX CC specification, but was obtained in electronic format directly from WIPO

XX CC at ftp.wipo.int/pub/published_pct_sequences.

XX CC

XX CC

XX CC

RESULT 13

AAAM31314
ID AAAM31314 standard; Protein; 66 AA.

AC AAAM31314;

DT 12-OCT-2001 (first entry)

DE Peptide #5234 encoded by probe for measuring cervical gene expression.

KM Probe; human; microarray; gene expression; cervical epithelial cell;

XX KW cervical cancer.

XX OS Homo sapiens.

PN WO200157278-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00670.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX PA Penn SG, Hanzel DK, Chen W, Rank DR;

XX PI WPI; 2001-488901/53.

XX DR WPI; 2001-488901/53.

XX PT Human genome-derived single exon nucleic acid probes useful for

XX PT analyzing gene expression in human cervical epithelial cells -

XX PS Claim 27; SEQ ID NO 23626; 487bp; English.

XX CC The present invention relates to human single exon nucleic acid probes

XX CC (SENP: see AA10068-AA128459). The present sequence is a peptide encoded

XX CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs

XX CC can be used to produce a single exon microarray, which can be used for

XX CC measuring human gene expression in a sample derived from human cervical

XX CC epithelial cells. By measuring gene expression, the probes are therefore

XX CC useful in grading and/or staging of diseases of the cervix, notably

XX CC cervical cancer.

XX CC Note: The sequence data for this patent did not form part of the printed

XX CC specification, but was obtained in electronic format directly from WIPO

XX CC at ftp.wipo.int/pub/published_pct_sequences.

XX CC

XX CC

XX CC

XX 17-OCT-2001 (first entry)
 XX Peptide #5351 encoded by probe for measuring placental gene expression.
 DE Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder.
 XX Homo sapiens.
 OS
 XX WO200157272-A2.
 PN
 XX 09-AUG-2001.
 PD
 XX 30-JAN-2001; 2001WO-US00663.
 PF
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-48897/53.
 DR
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -
 XX
 XX Claim 27; SEQ ID No 31583; 654bp; English.
 PS
 XX The present invention relates to single exon nucleic acid probes (SENPs;
 CC see A131315-A157546). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders.
 CC
 XX Sequence 66 AA;
 SQ
 Query Match 23.2%; Score 322; DB 22; Length 66;
 Best Local Similarity 100.0%; Pred. No. 3.5e-22;
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 103 TTGQALPATTETIOBASGPGADEVQVAPANALPARSEAAAVQPVIGISORVRMNSKEX 162
 DB 1 TTGQALPATTETIOBASGPGADEVQVAPANALPARSEAAAVQPVIGISORVRMNSKEX 60
 QY 163 KDLGTL 168
 DB 61 KDLGTL 66
 RESULT 14
 ABG40828
 ID ABG40828 standard; Peptide; 66 AA.
 XX
 AC ABG40828;
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE Human peptide encoded by genome-derived single exon probe SEQ ID 30493.
 XX
 KW Human: single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;

KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX
 OS Homo sapiens.
 XX
 XX WO200186003-A2.
 PN
 XX 15-NOV-2001.
 PD
 XX 30-JAN-2001; 2001WO-US00665.
 PF
 XX 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2002-114183/15.
 DR
 XX Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 XX
 XX Claim 27; SEQ ID No 30493; 634bp; English.
 PS
 XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes. The novel set of probes which hybridize at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridization of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridization to a single exon
 CC microarray having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from Wipo at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 66 AA;

Query Match 23.2%; Score 322; DB 23; Length 66;
 Best Local Similarity 100.0%; Pred. No. 3.5e-22;
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 TTSQLPAPFTTEIOBESGPGADEVQVFAPANMLPARSEMAAVQVPYIGISQFRMNSKER 162
 DB 1 TTSQLPAPFTTEIOBESGPGADEVQVFAPANMLPARSEMAAVQVPYIGISQFRMNSKER 60

QY 163 KDLGTL 168
 DB 61 KDLGTL 66

RESULT 15

AAV12615

ID AAV12615 standard; Protein: 56 AA.

AAV12615;

22-JUN-1999 (first entry)

Human 5' EST secreted protein SEQ ID NO: 280 from WO 9906553.

Human; secreted protein; EST; expressed sequence tag; diagnosis;
 forensic; gene therapy; chromosome mapping; signal peptide;
 upstream regulatory sequence; cytokine activity; cell proliferation;
 differentiation; haematopoiesis regulation; tissue growth regulation;
 reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
 thrombolytic; antiinflammatory; tumour inhibition; antitumour.

Homo sapiens.

WO9906553-A2.

11-FEB-1999.

31-JUL-1998; 98WO-IB01237.

01-AUG-1997; 97US-0905051.

(GEST) GENSET.

Duclet A, Dumas Mline Edwards J, Lacroix B;

MPI: 1999-153783/13.

N-PSDB; AAX41473.

New nucleic acids encoding human secreted proteins - obtained from
 cDNA libraries derived from umbilical cord, lymph ganglia,
 lymphocytes and placental tissue

Claim 34; Page 376; 411p; English.

AAX41379 to AAX41526 represent 5' expressed sequence tags (ESTs) for
 human secreted proteins, and encode the proteins given in AAV12521 to
 AAV12668, respectively. The proteins given represent the signal peptide
 and an N-terminal fragment of a secreted protein. The nucleic acid
 sequences can be used for producing secreted human gene products. They
 can also be used to develop products for diagnosis and therapy. The
 proteins obtained may have cytokine activity, cell
 proliferation/differentiation activity, haematopoiesis regulating
 activity, tissue growth regulating activity, reproductive hormone
 regulating activity, chemotactic/chemokinetic activity, haemostatic and
 thrombolytic activity, receptor/ligand activity, antiinflammatory
 activity, tumour inhibition activity or other activities. The products
 can be used in forensic, gene therapy and chromosome mapping procedures.
 The sequences can also be used for obtaining corresponding promoter
 sequences. The nucleic acids encoding the signal peptide can be used for
 directing extracellular secretion of a polypeptide or the insertion of a
 polypeptide into a membrane, or importing a polypeptide into a cell.

SQ Sequence + 56 AA;

Query Match 21.7%; Score 301; DB 20; Length 56;
 Best Local Similarity 98.2%; Pred. No. 2.4e-20;
 Matches 54; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIAVYQAEFLVSNMLIAEAYSGGCGFMDNGHLYREDQTSAPGIRCLINMLDAQSGL 55
 DB 2 LIAVYQAEFLVSNMLIAEAYSGGCGFMDNGHLYREDQTSAPGIRCLINMLDAQSGL 56

Search completed: April 7, 2003, 08:42:30

Job time : 38 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Comptgen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 08:43:33 ; Search time 16 Seconds

(without alignments)
1001.103 Million cell updates/sec

Title: US-10-057-951-2_COPY_2_263

Sequence: 1 LIAWVQAFIVSNMLLAAYG.....PVDPQSGRPLMGQACTGGA 262

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1388	100.0	263	9	US-10-102-704-2 Sequence 2, Appl1
2	1388	100.0	263	9	US-10-057-951-2 Sequence 2, Appl1
3	1388	100.0	263	10	US-09-084-491A-2 Sequence 2, Appl1
4	322	23.2	66	10	US-09-864-761-38457 Sequence 38457, A
5	154.5	11.1	527	9	US-09-987-457-18 Sequence 18, Appl1
6	154.5	11.1	527	9	US-09-987-455-19 Sequence 19, Appl1
7	154.5	11.1	562	9	US-09-974-298-145 Sequence 145, Appl1
8	154.5	11.1	562	10	US-09-969-271-7 Sequence 7, Appl1
9	142	10.2	160	9	US-09-335-325-35 Sequence 35, Appl1
10	142	10.2	160	9	US-10-131-241-35 Sequence 35, Appl1
11	142	10.2	160	10	US-09-761-120-35 Sequence 35, Appl1
12	142	10.2	250	9	US-09-335-325-30 Sequence 30, Appl1
13	142	10.2	250	10	US-10-131-241-30 Sequence 30, Appl1
14	142	10.2	250	10	US-09-761-120-30 Sequence 30, Appl1
15	142	10.2	260	9	US-10-131-241-61 Sequence 61, Appl1
16	142	10.2	339	9	US-09-335-325-3 Sequence 3, Appl1
17	142	10.2	339	9	US-10-131-241-3 Sequence 3, Appl1
18	142	10.2	339	10	US-09-761-120-3 Sequence 3, Appl1
19	142	10.2	339	10	US-09-761-120-3 Sequence 3, Appl1

20	142	10.2	352	9	US-09-335-325-40 Sequence 40, Appl1
21	142	10.2	352	9	US-10-131-241-40 Sequence 40, Appl1
22	142	10.2	352	10	US-09-761-120-40 Sequence 40, Appl1
23	142	10.2	368	10	US-09-761-120-42 Sequence 42, Appl1
24	142	10.2	378	9	US-09-335-325-42 Sequence 42, Appl1
25	142	10.2	378	9	US-10-131-241-42 Sequence 42, Appl1
26	142	10.2	378	10	US-09-873-676-1 Sequence 1, Appl1
27	142	10.2	458	10	US-09-946-893-4 Sequence 4, Appl1
28	142	10.2	569	10	US-09-946-893-5 Sequence 5, Appl1
29	142	10.2	571	10	US-09-946-893-8 Sequence 8, Appl1
30	142	10.2	576	10	US-09-946-893-6 Sequence 6, Appl1
31	142	10.2	791	9	US-09-967-386-1 Sequence 1, Appl1
32	142	10.2	810	10	US-09-946-893-2 Sequence 2, Appl1
33	138	9.9	79	9	US-09-335-325-8 Sequence 8, Appl1
34	138	9.9	79	9	US-10-131-241-8 Sequence 8, Appl1
35	138	9.9	79	10	US-09-753-064-2 Sequence 2, Appl1
36	138	9.9	79	10	US-09-761-120-8 Sequence 8, Appl1
37	135	9.7	377	9	US-09-887-455-8 Sequence 8, Appl1
38	133	9.6	326	9	US-10-057-951-3 Sequence 3, Appl1
39	133	9.6	354	9	US-09-987-457-10 Sequence 10, Appl1
40	133	9.6	354	9	US-09-987-455-11 Sequence 11, Appl1
41	133	9.6	372	9	US-10-102-704-3 Sequence 3, Appl1
42	133	9.6	372	10	US-09-084-491A-3 Sequence 3, Appl1
43	132	9.5	411	10	US-09-880-503-3 Sequence 3, Appl1
44	132	9.5	431	9	US-10-076-421-2 Sequence 2, Appl1
45	132	9.5	431	10	US-09-264-468B-1 Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-10-102-704-2
Sequence 2, Application US/10102704
Patent No US20020164768A1
GENERAL INFORMATION:
APPLICANT: Moore et al.
TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protein
FILE REFERENCE: PE378C1
CURRENT APPLICATION NUMBER: US/10/102,704
CURRENT FILING DATE: 2002-03-22
PRIOR APPLICATION NUMBER: 09/084,491
PRIOR FILING DATE: 1998-05-27
PRIOR APPLICATION NUMBER: 60/048,000
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
US-10-102-704-2

Query Match	Score 1388	DB 9	Length 263
Best Local Similarity	100.0%	Pred. No. 5,9e-119	
Matches 262	Conservative 0	Mismatches 0	Indels 0
Gaps 0			
QY	1	LIAWVQAFIVSNMLLAAYGSGGCFWNGHLYREDQTSAPAGLRCLNWDAGSLASAPV	60
DB	2	LIAWVQAFIVSNMLLAAYGSGGCFWNGHLYREDQTSAPAGLRCLNWDAGSLASAPV	61
QY	61	SGAGNSYCRNDEDEPRGWCVSGEAGVPEKRCCELRCPETTSQALPAFTTEIOEASE	120
DB	62	SGAGNSYCRNDEDEPRGWCVSGEAGVPEKRCCELRCPETTSQALPAFTTEIOEASE	121
QY	121	GGPADEVQFAPANLAPARSEAAAVPYTIGISQVRNNSKKKDLGTLGVVIGITMWTI	180
DB	122	GGPADEVQFAPANLAPARSEAAAVPYTIGISQVRNNSKKKDLGTLGVVIGITMWTI	181
QY	181	IAAGAGIIGYGYKXGKDLKEQHDQVCEREMQITPLSAFTNPCEIIVDEKTVVHTS	240
DB	182	IAAGAGIIGYGYKXGKDLKEQHDQVCEREMQITPLSAFTNPCEIIVDEKTVVHTS	241

```
QY 241 QTPVDPQEGSTPLMGQAGTPGA 262
|||
DB 242 QTPVDPQEGSTPLMGQAGTPGA 263

RESULT 2
US-10-057-951-2
Sequence 2, Application US/10057951
Patent No. US20020177213A1
GENERAL INFORMATION:
APPLICANT: Moore et al.
TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
FILE REFERENCE: PF378P1
CURRENT APPLICATION NUMBER: US/10/057,951
PRIOR FILING DATE: 2002-01-29
PRIOR APPLICATION NUMBER: US 09/411,977
PRIOR FILING DATE: 1999-10-04
PRIOR APPLICATION NUMBER: US 09/084,491
PRIOR FILING DATE: 1998-05-27
PRIOR APPLICATION NUMBER: US 60/048,000
PRIOR FILING DATE: 1997-05-28
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 2
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
US-10-057-951-2

Query Match
Best Local Similarity 100.0%; Score 1388; DB 9; Length 263;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLAWQAFVSNMLLAEAAYSGGCFMNDGHLVREDQTSPPAGLRCLNMLDAQSGLASAPV 60
DB 2 LLAWQAFVSNMLLAEAAYSGGCFMNDGHLVREDQTSPPAGLRCLNMLDAQSGLASAPV 61
QY 61 SGAGNSHYCRNPDEDPKRCVSGEAGVPEKRCPCEDLRCPETTSQALPAFTTEIOEASE 120
DB 62 SGAGNSHYCRNPDEDPKRCVSGEAGVPEKRCPCEDLRCPETTSQALPAFTTEIOEASE 121
QY 121 GPGADEVQVAPANALPARSEAAVQPIVIGISQVRNNSKEKDLGTLGVIGITMVI 180
DB 122 GPGADEVQVAPANALPARSEAAVQPIVIGISQVRNNSKEKDLGTLGVIGITMVI 181
QY 181 IAIAGIILGYSYKRGDKLKEQHDQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTS 240
DB 182 IAIAGIILGYSYKRGDKLKEQHDQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTS 241
QY 241 QTPVDPQEGSTPLMGQAGTPGA 262
DB 242 QTPVDPQEGSTPLMGQAGTPGA 263

RESULT 3
US-09-084-491A-2
Sequence 2, Application US/09084491A
Patent No. US20020061576A1
GENERAL INFORMATION:
APPLICANT: MOORE, PAUL A.
APPLICANT: RUBEN, STEVEN M.
APPLICANT: REBER, REINHARD
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
```

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/084,491A
FILING DATE: 27-MAY-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BROOKS, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF378
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-084-491A-2

Query Match
Best Local Similarity 100.0%; Score 1388; DB 10; Length 263;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLAWQAFVSNMLLAEAAYSGGCFMNDGHLVREDQTSPPAGLRCLNMLDAQSGLASAPV 60
DB 2 LLAWQAFVSNMLLAEAAYSGGCFMNDGHLVREDQTSPPAGLRCLNMLDAQSGLASAPV 61
QY 61 SGAGNSHYCRNPDEDPKRCVSGEAGVPEKRCPCEDLRCPETTSQALPAFTTEIOEASE 120
DB 62 SGAGNSHYCRNPDEDPKRCVSGEAGVPEKRCPCEDLRCPETTSQALPAFTTEIOEASE 121
QY 121 GPGADEVQVAPANALPARSEAAVQPIVIGISQVRNNSKEKDLGTLGVIGITMVI 180
DB 122 GPGADEVQVAPANALPARSEAAVQPIVIGISQVRNNSKEKDLGTLGVIGITMVI 181
QY 181 IAIAGIILGYSYKRGDKLKEQHDQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTS 240
DB 182 IAIAGIILGYSYKRGDKLKEQHDQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTS 241
QY 241 QTPVDPQEGSTPLMGQAGTPGA 262
DB 242 QTPVDPQEGSTPLMGQAGTPGA 263

RESULT 4
US-09-864-761-38457
Sequence 38457, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38457
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC002073.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
; OTHER INFORMATION: SWISSPROT HIT: P39881, EVALU8 3.40e+00
; OTHER INFORMATION: EST_HUMAN HIT: BE299406.1, EVALU8 5.00e-30
US-09-864-761-38457
```

```
Query Match 23.2%; Score 322; DB 10; Length 66;
Best Local Similarity 100.0%; Pred. No. 1.7e-22;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 103 TTGQALPAFTTETQASSEGPADEVQVAPANALPARSEAAAVQVIGISQRYMNSKEK 162
Db 1 TTGQALPAFTTETQASSEGPADEVQVAPANALPARSEAAAVQVIGISQRYMNSKEK 60
```

```
QY 163 KDLGTL 168
Db 61 KDLGTL 66
```

```
RESULT 5
US-09-987-457-18
; Sequence 18, Application US/09987457
; Publication No. US20030013150A1
; GENERAL INFORMATION:
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Uiradej
; APPLICANT: Tayapiwatana, Chatchai
; APPLICANT: Goetz, Friedrich
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652.2180001
; CURRENT APPLICATION NUMBER: US/09/987,457
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,573
; PRIOR FILING DATE: 2001-02-15
```

```
; PRIOR APPLICATION NUMBER: GB 00 27 782.2
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens (tpa)
US-09-987-457-18
```

```
Query Match 11.1%; Score 154.5; DB 9; Length 527;
Best Local Similarity 39.6%; Pred. No. 4.7e-06;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
```

```
QY 24 CFMDNCHLYREDQTSFAPGLACLMWLDQSLASAPVS-----GAGNHSYCRNPDE 74
Db 92 CYBDQGISIRGTWSTAESGACETW--NSSALQKPYSGRRPDALRLGNNHYCRNPDR 149
QY 75 DPRGPMCVYSGEAGVPEKRPCEDLRCPETTS 105
Db 150 DSK-FWCYVF-KAGKYSSEFCSTPACSEGNS 178
```

```
RESULT 6
US-09-987-455-19
; Sequence 19, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Uiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Goetz, Friedrich
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for large Scale Production of Recombinant
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-987-455-19
```

```
Query Match 11.1%; Score 154.5; DB 9; Length 527;
Best Local Similarity 39.6%; Pred. No. 4.7e-06;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
```

```
QY 24 CFMDNCHLYREDQTSFAPGLACLMWLDQSLASAPVS-----GAGNHSYCRNPDE 74
Db 92 CYBDQGISIRGTWSTAESGACETW--NSSALQKPYSGRRPDALRLGNNHYCRNPDR 149
QY 75 DPRGPMCVYSGEAGVPEKRPCEDLRCPETTS 105
Db 150 DSK-FWCYVF-KAGKYSSEFCSTPACSEGNS 178
```

```
RESULT 7
US-09-974-298-145
; Sequence 145, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hui-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
```

```

; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 145
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 1001470CD1
US-09-974-298-145
```

```

Query Match 11.1%; Score 154.5; DB 9; Length 562;
Best Local Similarity 39.6%; Pred. No. 5.1e-06;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
```

```

Qy 24 CFWDNGHLYREDQTSBPAGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 74
Db 127 CYEDQGISYRGWTSTASGAECTNW--NSSALQKPYSGRRPDATRLGIGHNHYCRNPDR 184
```

```

Qy 75 DPRGWCYVSGEAGVPEKRPCEDLRCPEETS 105
Db 185 DSK-PWCYTF-KAGKYSSEFCSTPACSEGN 213
```

```

RESULT 8
US-09-969-271-7
; Sequence 7, Application US/09969271
; Patent No. US20020098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT FILING DATE: 2001-10-01
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: GB 0025473.0
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-969-271-7
```

```

Query Match 11.1%; Score 154.5; DB 10; Length 562;
Best Local Similarity 39.6%; Pred. No. 5.1e-06;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
```

```

Qy 24 CFWDNGHLYREDQTSBPAGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 74
Db 127 CYEDQGISYRGWTSTASGAECTNW--NSSALQKPYSGRRPDATRLGIGHNHYCRNPDR 184
```

```

Qy 75 DPRGWCYVSGEAGVPEKRPCEDLRCPEETS 105
Db 185 DSK-PWCYTF-KAGKYSSEFCSTPACSEGN 213
```

```

RESULT 9
US-09-335-325-35
; Sequence 35, Application US/09335325
; Patent No. US20020164717A1
; GENERAL INFORMATION:
; APPLICANT: Polman, M. Judah
; O'Reilly, Michael
; Cao, Yihai
; Sim, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESS: Jones & Askew
```

```

; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 160 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: KI-2
; SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-335-325-35
```

```

Query Match 10.2%; Score 142; DB 9; Length 160;
Best Local Similarity 37.2%; Pred. No. 1.4e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;
```

```

Qy 24 CFWDNGHLYREDQTSBPAGRLCLNWLDA-----OSGLASAPVSGAGNHSYCRNPDEDPG 78
Db 1 CKTNGKNGVGTWSTKRNKIGITCKKWSSTSPHRRPRSPATHSEGL-BENYCRNPDPDPOG 59
```

```

Qy 79 PWCYVSGEAGVPEKRPCEDLRCPE 102
Db 60 PWCYTTD---PEKRYDYCDILECEE 81
```

```

RESULT 10
US-10-131-241-35
; Sequence 35, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT FILING DATE: 2002-07-22
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
```

```
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 35
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-241-35

Query Match      10.2%; Score 142; DB 9; Length 160;
Best Local Similarity 37.2%; Pred. No. 1.4e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

QY 24 CFMDNGHLYREDQTSAPAGRLCLNWLDA-----QSGLASAPVSGAGNHSYCRNPDEPRG 78
DB 1 CKTGNGKXVIGTWSKTKNGITCKQKWSSTSPHRRPFSPTHSEGL-ENYCNPNDDPQG 59
QY 79 PWCYVSGEAGVPEKR--PCEDLRCE 102
DB 60 PWCYTTD---PEKRYDYCDILECEE 81

RESULT 11
US-09-761-120-35
; Sequence 35, Application US/09761120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasminogen
; FILE REFERENCE: 05940-0151 (43171-252068)
; CURRENT APPLICATION NUMBER: US/09/761,120
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/309,821
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 08/866,735
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 35
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Kringle 1-2
US-09-761-120-35

Query Match      10.2%; Score 142; DB 10; Length 160;
Best Local Similarity 37.2%; Pred. No. 1.4e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

QY 24 CFMDNGHLYREDQTSAPAGRLCLNWLDA-----QSGLASAPVSGAGNHSYCRNPDEPRG 78
DB 1 CKTGNGKXVIGTWSKTKNGITCKQKWSSTSPHRRPFSPTHSEGL-ENYCNPNDDPQG 59
QY 79 PWCYVSGEAGVPEKR--PCEDLRCE 102
DB 60 PWCYTTD---PEKRYDYCDILECEE 81

RESULT 12
US-09-335-325-30
; Sequence 30, Application US/09335325
; Patent No. US20020164717A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; O'Reilly, Michael
; Cao, Yihai
; Sain, B. Kim Lee
; TITLE OF INVENTION: Angiostatin Fragments and Method of Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
```

```
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/335,325
; FILING DATE: 17-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,788
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: K1-3
; SEQUENCE DESCRIPTION: SEQ ID NO: 30:
US-09-335-325-30

Query Match      10.2%; Score 142; DB 9; Length 250;
Best Local Similarity 37.2%; Pred. No. 2.5e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

QY 24 CFMDNGHLYREDQTSAPAGRLCLNWLDA-----QSGLASAPVSGAGNHSYCRNPDEPRG 78
DB 1 CKTGNGKXVIGTWSKTKNGITCKQKWSSTSPHRRPFSPTHSEGL-ENYCNPNDDPQG 59
QY 79 PWCYVSGEAGVPEKR--PCEDLRCE 102
DB 60 PWCYTTD---PEKRYDYCDILECEE 81

RESULT 13
US-10-131-241-30
; Sequence 30, Application US/10131241
; Patent No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patentin version 3.1
```

Mon Apr 7 10:22:35 2003

us-10-057-951-2_copy_2_263.rapb

Page 6

SEQ ID NO 30
LENGTH: 250
TYPE: PRT
ORGANISM: Homo sapiens
US-10-131-241-30

Query Match 10.2%; Score 142; DB 9; Length 250;
Best Local Similarity 37.2%; Pred. No. 2.5e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

QY 24 CFWDNGHLYREDQTSAPGLRCLNWLDA-----QSLASAPVSGAGNHSYCRNPDEDPFG 78
DB 1 CKTGNGKXVYRGTMKTKNGITTCOKMSSTPHPRFSPATHPSGL-ENYCRNPNDPFG 59
QY 79 PWCYVSGEAGVPEKR--PCEDLRCP 102
DB 60 PWCYTTD---PEKRYDYCDILECEE 81

RESULT 14
US-09-761-120-30
Sequence 30 Application US/09761120
Patent No US20020037847A1
GENERAL INFORMATION:
APPLICANT: Folman, M. Judah
APPLICANT: O'Reilly, Michael
TITLE OF INVENTION: Nucleic Acid Encoding Kringle 1-5 Region Fragments of Plasminogen
FILE REFERENCE: 05940-0151 (43171-252068)
CURRENT APPLICATION NUMBER: US/09/761,120
PRIOR APPLICATION NUMBER: 2001-01-16
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 08/866,735
PRIOR FILING DATE: 1997-05-30
NUMBER OF SEQ ID NOS: 47
SOFTWARE: PatentIn version 3.0
SEQ ID NO 30
LENGTH: 250
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Kringle 1-3
US-09-761-120-30

Query Match 10.2%; Score 142; DB 10; Length 250;
Best Local Similarity 37.2%; Pred. No. 2.5e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

QY 24 CFWDNGHLYREDQTSAPGLRCLNWLDA-----QSLASAPVSGAGNHSYCRNPDEDPFG 78
DB 1 CKTGNGKXVYRGTMKTKNGITTCOKMSSTPHPRFSPATHPSGL-ENYCRNPNDPFG 59
QY 79 PWCYVSGEAGVPEKR--PCEDLRCP 102
DB 60 PWCYTTD---PEKRYDYCDILECEE 81

RESULT 15
US-10-131-241-61
Sequence 61 Application US/10131241
Publication No. US20030012792A1
GENERAL INFORMATION:
APPLICANT: Holaday, John W.
APPLICANT: Fortier, Anne H.
TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
FILE REFERENCE: 05213-0344 43170-271565
CURRENT APPLICATION NUMBER: US/10/131,241
CURRENT FILING DATE: 2002-07-22
PRIOR APPLICATION NUMBER: US 09/413,049
PRIOR FILING DATE: 1999-10-06
PRIOR APPLICATION NUMBER: US 09/316,802

PRIOR FILING DATE: 1999-05-21
PRIOR APPLICATION NUMBER: US 60/086,586
PRIOR FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.1
SEQ ID NO 61
LENGTH: 260
TYPE: PRT
ORGANISM: Homo sapiens
US-10-131-241-61

Query Match 10.2%; Score 142; DB 9; Length 260;
Best Local Similarity 37.2%; Pred. No. 2.6e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

QY 24 CFWDNGHLYREDQTSAPGLRCLNWLDA-----QSLASAPVSGAGNHSYCRNPDEDPFG 78
DB 6 CKTGNGKXVYRGTMKTKNGITTCOKMSSTPHPRFSPATHPSGL-ENYCRNPNDPFG 64
QY 79 PWCYVSGEAGVPEKR--PCEDLRCP 102
DB 65 PWCYTTD---PEKRYDYCDILECEE 86

Search completed: April 7, 2003, 08:50:49
Job time: 17 secs

GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 08:42:37 ; Search time 328 Seconds

(without alignments)
515.000 Million cell updates/sec

Title: US-10-057-951-2_COPY_2_263

Perfect score: 1388
Sequence: 1 LLAWQAFIVSNMLLAAYG.....PVDPEGSTPLMGQAPGA 262Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending Patents AA Main:*

1: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US099_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US103_COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US104_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1388	100.0	263	1	PCT-US00-27239-2
2	1388	100.0	263	1	PCT-US02-05301-164
3	1388	100.0	263	1	PCT-US98-10728-2
4	1388	100.0	263	14	US-09-084-9314-2
5	1388	100.0	263	15	US-09-144-889-2
6	1388	100.0	263	24	US-10-057-951-2

7	1388	100.0	263	25	US-10-102-704-2	Sequence 2, Appl1
8	1388	100.0	263	27	US-60-048-000-2	Sequence 2, Appl1
9	1385	99.8	263	21	US-09-791-537-39562	Sequence 39562, A
10	1385	99.8	263	23	US-09-927-796-44	Sequence 44, Appl1
11	1385	99.8	263	26	US-10-210-951-44	Sequence 44, Appl1
12	1385	99.8	263	26	US-10-211-856-44	Sequence 44, Appl1
13	1385	99.8	263	26	US-10-211-884-44	Sequence 44, Appl1
14	1385	99.8	263	27	US-60-230-435-1590	Sequence 1590, Ap
15	1385	99.8	263	27	US-60-389-987-244	Sequence 244, App
16	1385	99.8	263	27	US-60-412-418-244	Sequence 244, App
17	1374	99.0	263	20	US-09-611-526-3727	Sequence 3727, Ap
18	1334	96.1	257	14	PCT-US01-14827-8898	Sequence 8898, Ap
19	1328	95.7	250	14	US-09-001-403-64	Sequence 64, Appl
20	1320.5	95.1	295	1	PCT-US01-14827-8900	Sequence 8900, Ap
21	1265.5	91.2	308	27	US-60-207-315-471	Sequence 471, App
22	1162.5	83.8	178	15	US-09-144-889-4	Sequence 4, Appl1
23	948	68.3	186	27	US-60-213-800-237	Sequence 237, App
24	747	53.8	146	1	PCT-US02-05301-237	Sequence 237, App
25	747	53.8	146	1	US-09-307-140-979	Sequence 311, App
26	695	50.1	126	17	US-09-817-076-979	Sequence 979, App
27	695	50.1	126	22	US-09-817-076-979	Sequence 979, App
28	607	43.7	109	22	US-09-817-076-979	Sequence 979, App
29	607	43.7	109	27	US-60-197-873-13839	Sequence 13839, A
30	575	41.4	117	1	PCT-US01-14827-8897	Sequence 8897, Ap
31	426	30.7	88	22	US-09-834-866-16200	Sequence 16200, A
32	426	30.7	88	27	US-60-197-873-16200	Sequence 16200, A
33	322	23.2	66	1	PCT-US01-00663-31583	Sequence 31583, A
34	322	23.2	66	22	US-09-864-761-38457	Sequence 38457, A
35	322	23.2	66	25	US-10-182-993-30642	Sequence 30642, A
36	322	23.2	66	25	US-10-182-993-30642	Sequence 30642, A
37	322	23.2	66	25	US-10-182-993-24629	Sequence 24629, A
38	322	23.2	66	25	US-10-182-993-24629	Sequence 24629, A
39	322	23.2	66	26	US-10-203-134-3143	Sequence 3143, A
40	322	23.2	66	26	US-10-203-135-30493	Sequence 30493, A
41	322	23.2	66	26	US-10-203-136-31400	Sequence 31400, A
42	322	23.2	66	26	US-10-203-137-31583	Sequence 31583, A
43	301	21.7	56	26	US-10-203-139-30540	Sequence 30540, A
44	301	21.7	56	13	US-09-905-051-280	Sequence 280, App
45	300	21.6	55	13	US-08-905-135-428	Sequence 428, App

ALIGNMENTS

RESULT 1
PCT-US00-27239-2
Sequence 2, Application PC/TUS0027239
GENERAL INFORMATION:
APPLICANT: HUMAN GENOME SCIENCES, INC.
TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
FILE REFERENCE: P378PCT2
CURRENT APPLICATION NUMBER: PCT/US00/27239
CURRENT FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 09/411,977
PRIOR FILING DATE: 1999-10-04
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US00-27239-2

Query Match 100.0%; Score 1388; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.8e-128;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLAWQAFIVSNMLLAAYGSGGCFWMDNGHLYREQOTSPAPGLCLNWLDAOSGLSARV 60
DB 2 LLAWQAFIVSNMLLAAYGSGGCFWMDNGHLYREQOTSPAPGLCLNWLDAOSGLSARV 61
QY 61 SGGGHSYGRNDEDPGPFCTVSGAGVPEKRCPECDRCGETTSQALPFTETIOESE 120

```
Db 62 SGAGNHSYCRNPDEDPGPMCVYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASE 121
Qy 121 GGADEVQVFPANALPARSEAAAVQVIGISQVRNMSKEKDLGLTGLVIGITMVYI 180
Db 122 GGADEVQVFPANALPARSEAAAVQVIGISQVRNMSKEKDLGLTGLVIGITMVYI 181
Qy 181 IAIAGGIIIGSYKRGKDLKEQHDQVCEREMQRTITPLSAFTNPCEIYDEKTVVHTS 240
Db 182 IAIAGGIIIGSYKRGKDLKEQHDQVCEREMQRTITPLSAFTNPCEIYDEKTVVHTS 241
Qy 241 QTPVPDQEGSTPLMGOAGTGA 262
Db 242 QTPVPDQEGSTPLMGOAGTGA 263

RESULT 2
PCT-US92-05301-164
; Sequence 164, Application PC/TUS0205301
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: PS736PCT
; CURRENT APPLICATION NUMBER: PCT/US02/05301
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/304,417
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US 60/270,625
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 164
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-05301-164

Query Match 100.0%; Score 1388; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 1,8e-128;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIAWQAFIVSNMLAEAYSGGCFWMDNGHLTYREDQTSAPAGLRCLNMLDAOSGLASAPV 60
Db 2 LIAWQAFIVSNMLAEAYSGGCFWMDNGHLTYREDQTSAPAGLRCLNMLDAOSGLASAPV 61
Qy 61 SGAGNHSYCRNPDEDPGPMCVYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASE 120
Db 62 SGAGNHSYCRNPDEDPGPMCVYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASE 121
Qy 121 GGADEVQVFPANALPARSEAAAVQVIGISQVRNMSKEKDLGLTGLVIGITMVYI 180
Db 122 GGADEVQVFPANALPARSEAAAVQVIGISQVRNMSKEKDLGLTGLVIGITMVYI 181
Qy 181 IAIAGGIIIGSYKRGKDLKEQHDQVCEREMQRTITPLSAFTNPCEIYDEKTVVHTS 240
Db 182 IAIAGGIIIGSYKRGKDLKEQHDQVCEREMQRTITPLSAFTNPCEIYDEKTVVHTS 241
Qy 241 QTPVPDQEGSTPLMGOAGTGA 262
Db 242 QTPVPDQEGSTPLMGOAGTGA 263

RESULT 3
PCT-US98-10728-2
; Sequence 2, Application PC/TUS9810728
; GENERAL INFORMATION:
; APPLICANT: BENNER, REINHARD
; APPLICANT: MOORE, PAUL
; APPLICANT: RUBEN, STEVEN
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
```

```
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/10728
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A.
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF379PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US98-10728-2

Query Match 100.0%; Score 1388; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 1,8e-128;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIAWQAFIVSNMLAEAYSGGCFWMDNGHLTYREDQTSAPAGLRCLNMLDAOSGLASAPV 60
Db 2 LIAWQAFIVSNMLAEAYSGGCFWMDNGHLTYREDQTSAPAGLRCLNMLDAOSGLASAPV 61
Qy 61 SGAGNHSYCRNPDEDPGPMCVYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASE 120
Db 62 SGAGNHSYCRNPDEDPGPMCVYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASE 121
Qy 121 GGADEVQVFPANALPARSEAAAVQVIGISQVRNMSKEKDLGLTGLVIGITMVYI 180
Db 122 GGADEVQVFPANALPARSEAAAVQVIGISQVRNMSKEKDLGLTGLVIGITMVYI 181
Qy 181 IAIAGGIIIGSYKRGKDLKEQHDQVCEREMQRTITPLSAFTNPCEIYDEKTVVHTS 240
Db 182 IAIAGGIIIGSYKRGKDLKEQHDQVCEREMQRTITPLSAFTNPCEIYDEKTVVHTS 241
Qy 241 QTPVPDQEGSTPLMGOAGTGA 262
Db 242 QTPVPDQEGSTPLMGOAGTGA 263

RESULT 4
US-09-084-491A-2
; Sequence 2, Application US/09084491A
; GENERAL INFORMATION:
; APPLICANT: MOORE, PAUL A.
; APPLICANT: RUBEN, STEVEN M.
; APPLICANT: BENNER, REINHARD
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

SOFTWARE: Patent In Release #1.0, Version #1.10
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/084,491A
FILING DATE: 27-MAY-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF378
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-084-491A-2

Query Match 100.0%; Score 1388; DB 14; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.8e-128;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLAMVOAFIVSNMLAEAYSGGCFWMDNGHLTYREDQTSAPAGLRCLNWLDAOSGLASAPV 60
Db 2 LLAMVOAFIVSNMLAEAYSGGCFWMDNGHLTYREDQTSAPAGLRCLNWLDAOSGLASAPV 61
Qy 61 SGAGNHSYCRNPDEDPGPMCVYSGAGVPEKRPCEDLRCPEETTSQALPAFTTEIOEASE 120
Db 62 SGAGNHSYCRNPDEDPGPMCVYSGAGVPEKRPCEDLRCPEETTSQALPAFTTEIOEASE 121
Qy 121 GPADDEVQVAPANALPARSEAAAVQPVIGISQVRMNSKEKDLGTLGYVLGITMVI 180
Db 122 GPADDEVQVAPANALPARSEAAAVQPVIGISQVRMNSKEKDLGTLGYVLGITMVI 181
Qy 181 IAIAGIILIGSYKRGKDLKEQDQKVCEREMORTLPLSAFTNPTCEIYDEKTVVHTS 240
Db 182 IAIAGIILIGSYKRGKDLKEQDQKVCEREMORTLPLSAFTNPTCEIYDEKTVVHTS 241
Qy 241 QTPVDPQSGSTPLMGOAGTPGA 262
Db 242 QTPVDPQSGSTPLMGOAGTPGA 263

RESULT 5
US-09-144-889-2

; Sequence 2, Application US/09144889B
; GENERAL INFORMATION:
; APPLICANT: Earl F. Albone
; APPLICANT: Kristine K. Kikly
; TITLE OF INVENTION: KRINGLE1
; FILE REFERENCE: GH-70249
; CURRENT APPLICATION NUMBER: US/09/144,889B
; CURRENT FILING DATE: 1998-09-01
; EARLIER APPLICATION NUMBER: 60/056,032
; EARLIER FILING DATE: 1997-09-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-144-889-2

Query Match 100.0%; Score 1388; DB 15; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.8e-128;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLAMVOAFIVSNMLAEAYSGGCFWMDNGHLTYREDQTSAPAGLRCLNWLDAOSGLASAPV 60
Db 2 LLAMVOAFIVSNMLAEAYSGGCFWMDNGHLTYREDQTSAPAGLRCLNWLDAOSGLASAPV 61

Qy 61 SGAGNHSYCRNPDEDPGPMCVYSGAGVPEKRPCEDLRCPEETTSQALPAFTTEIOEASE 120
Db 62 SGAGNHSYCRNPDEDPGPMCVYSGAGVPEKRPCEDLRCPEETTSQALPAFTTEIOEASE 121
Qy 121 GPADDEVQVAPANALPARSEAAAVQPVIGISQVRMNSKEKDLGTLGYVLGITMVI 180
Db 122 GPADDEVQVAPANALPARSEAAAVQPVIGISQVRMNSKEKDLGTLGYVLGITMVI 181
Qy 181 IAIAGIILIGSYKRGKDLKEQDQKVCEREMORTLPLSAFTNPTCEIYDEKTVVHTS 240
Db 182 IAIAGIILIGSYKRGKDLKEQDQKVCEREMORTLPLSAFTNPTCEIYDEKTVVHTS 241
Qy 241 QTPVDPQSGSTPLMGOAGTPGA 262
Db 242 QTPVDPQSGSTPLMGOAGTPGA 263

RESULT 6
US-10-057-951-2

; Sequence 2, Application US/10057951
; GENERAL INFORMATION:
; APPLICANT: Moore et al.
; TITLE OF INVENTION: Tissue Plasminogen Activator-like Protease
; FILE REFERENCE: PF378P1
; CURRENT APPLICATION NUMBER: US/10/057,951
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US 09/411,977
; PRIOR FILING DATE: 1999-10-04
; PRIOR APPLICATION NUMBER: US 09/084,491
; PRIOR FILING DATE: 1998-05-27
; PRIOR APPLICATION NUMBER: US 60/048,000
; PRIOR FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent In Ver. 3.1
; SEQ ID NO: 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-951-2

Query Match 100.0%; Score 1388; DB 24; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.8e-128;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLAMVOAFIVSNMLAEAYSGGCFWMDNGHLTYREDQTSAPAGLRCLNWLDAOSGLASAPV 60
Db 2 LLAMVOAFIVSNMLAEAYSGGCFWMDNGHLTYREDQTSAPAGLRCLNWLDAOSGLASAPV 61
Qy 61 SGAGNHSYCRNPDEDPGPMCVYSGAGVPEKRPCEDLRCPEETTSQALPAFTTEIOEASE 120
Db 62 SGAGNHSYCRNPDEDPGPMCVYSGAGVPEKRPCEDLRCPEETTSQALPAFTTEIOEASE 121
Qy 121 GPADDEVQVAPANALPARSEAAAVQPVIGISQVRMNSKEKDLGTLGYVLGITMVI 180
Db 122 GPADDEVQVAPANALPARSEAAAVQPVIGISQVRMNSKEKDLGTLGYVLGITMVI 181
Qy 181 IAIAGIILIGSYKRGKDLKEQDQKVCEREMORTLPLSAFTNPTCEIYDEKTVVHTS 240
Db 182 IAIAGIILIGSYKRGKDLKEQDQKVCEREMORTLPLSAFTNPTCEIYDEKTVVHTS 241
Qy 241 QTPVDPQSGSTPLMGOAGTPGA 262
Db 242 QTPVDPQSGSTPLMGOAGTPGA 263

RESULT 7
US-10-102-704-2

; Sequence 2, Application US/10102704
; GENERAL INFORMATION:
; APPLICANT: Moore et al.
; TITLE OF INVENTION: Tissue Plasminogen Activator-like Protein
; FILE REFERENCE: PF378C1
; CURRENT APPLICATION NUMBER: US/10/102,704

```

; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/084,491
; PRIOR FILING DATE: 1998-05-27
; PRIOR APPLICATION NUMBER: 60/049,000
; PRIOR FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent version 3.1
; SEQ ID NO: 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-102-704-2

Query Match
Best Local Similarity 100.0%; Score 1388; DB 25; Length 263;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLAMVOAFLVSNMLLAAYSGGCFWMDNGHLYREDQTSPPAGLRCLNMLDQSGIASAPV 60
DB 2 LLAMVOAFLVSNMLLAAYSGGCFWMDNGHLYREDQTSPPAGLRCLNMLDQSGIASAPV 61
QY 61 SGAGNHSYCRNPDDPPGPMCYSGAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASE 120
DB 62 SGAGNHSYCRNPDDPPGPMCYSGAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASE 121
QY 121 GGADEVGVFAPANAALPARSEAAAVOPVIGISQVRNNSKEKKDLGTIGYVLGITMVI 180
DB 122 GGADEVGVFAPANAALPARSEAAAVOPVIGISQVRNNSKEKKDLGTIGYVLGITMVI 181
QY 181 IAIAGAILIGSYRGKDLKEQHDQVCEREMORITPLSAFTNPCEIVDEKTVVHTS 240
DB 182 IAIAGAILIGSYRGKDLKEQHDQVCEREMORITPLSAFTNPCEIVDEKTVVHTS 241
QY 241 QTEVDPQSGSTPLMGQAGTPGA 262
DB 242 QTEVDPQSGSTPLMGQAGTPGA 263

RESULT 8
US-60-048-000-2
; Sequence 2, Application US/60048000
; GENERAL INFORMATION:
; APPLICANT: EBER, REINHARD
; APPLICANT: MOORE, PAUL
; APPLICANT: RUBEN, STEVE
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/048,000
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A.
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF379PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 263 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-60-048-000-2

Query Match
Best Local Similarity 100.0%; Score 1388; DB 27; Length 263;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLAMVOAFLVSNMLLAAYSGGCFWMDNGHLYREDQTSPPAGLRCLNMLDQSGIASAPV 60
DB 2 LLAMVOAFLVSNMLLAAYSGGCFWMDNGHLYREDQTSPPAGLRCLNMLDQSGIASAPV 61
QY 61 SGAGNHSYCRNPDDPPGPMCYSGAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASE 120
DB 62 SGAGNHSYCRNPDDPPGPMCYSGAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASE 121
QY 121 GGADEVGVFAPANAALPARSEAAAVOPVIGISQVRNNSKEKKDLGTIGYVLGITMVI 180
DB 122 GGADEVGVFAPANAALPARSEAAAVOPVIGISQVRNNSKEKKDLGTIGYVLGITMVI 181
QY 181 IAIAGAILIGSYRGKDLKEQHDQVCEREMORITPLSAFTNPCEIVDEKTVVHTS 240
DB 182 IAIAGAILIGSYRGKDLKEQHDQVCEREMORITPLSAFTNPCEIVDEKTVVHTS 241
QY 241 QTEVDPQSGSTPLMGQAGTPGA 262
DB 242 QTEVDPQSGSTPLMGQAGTPGA 263

RESULT 9
US-09-791-537-39562
; Sequence 39562, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomomix, Inc.
; APPLICANT: Dede, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB.
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent version 3.0
; SEQ ID NO: 39562
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-39562

Query Match
Best Local Similarity 99.6%; Score 1385; DB 21; Length 263;
Matches 261; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLAMVOAFLVSNMLLAAYSGGCFWMDNGHLYREDQTSPPAGLRCLNMLDQSGIASAPV 60
DB 2 LLAMVOAFLVSNMLLAAYSGGCFWMDNGHLYREDQTSPPAGLRCLNMLDQSGIASAPV 61
QY 61 SGAGNHSYCRNPDDPPGPMCYSGAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASE 120
DB 62 SGAGNHSYCRNPDDPPGPMCYSGAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASE 121
QY 121 GGADEVGVFAPANAALPARSEAAAVOPVIGISQVRNNSKEKKDLGTIGYVLGITMVI 180
DB 122 GGADEVGVFAPANAALPARSEAAAVOPVIGISQVRNNSKEKKDLGTIGYVLGITMVI 181
QY 181 IAIAGAILIGSYRGKDLKEQHDQVCEREMORITPLSAFTNPCEIVDEKTVVHTS 240
DB 182 IAIAGAILIGSYRGKDLKEQHDQVCEREMORITPLSAFTNPCEIVDEKTVVHTS 241
QY 241 QTEVDPQSGSTPLMGQAGTPGA 262
DB 242 QTEVDPQSGSTPLMGQAGTPGA 263

```

Db 242 @TRVDPQECTPLMGQATPBA 263

```

RESULT 10
US-09-927-796-44
Sequence 44, Application US/09927796
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Guiney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Masters, Scot A.
APPLICANT: Pan, James
APPLICANT: Pitti, Robert M.
APPLICANT: Roy, Margaret Ann
APPLICANT: Smith, Victoria
APPLICANT: Stone, Donna M.
APPLICANT: Matenabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
FILE REFERENCE: P2331R1C1
CURRENT APPLICATION NUMBER: US/09/927,796
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 60/014699
PRIOR FILING DATE: 1996-04-01
PRIOR APPLICATION NUMBER: 60/026943
PRIOR FILING DATE: 1996-09-23
PRIOR APPLICATION NUMBER: 60/059121
PRIOR FILING DATE: 1997-07-17
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/062037
PRIOR FILING DATE: 1997-10-10
PRIOR APPLICATION NUMBER: 60/063755
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063045
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063046
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/065511
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066772
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/067411
PRIOR FILING DATE: 1997-12-03
PRIOR APPLICATION NUMBER: 60/069862
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/095929
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/097978
PRIOR FILING DATE: 1998-08-26
PRIOR APPLICATION NUMBER: 60/103396
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: 60/108867
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/112851
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/119965
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: 60/123972
PRIOR FILING DATE: 1999-03-11
PRIOR APPLICATION NUMBER: 60/133459
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 60/140650
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/140653
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/144758
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698

```

```

PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/146222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: 60/149395
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/151689
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/625328
PRIOR FILING DATE: 1996-04-01
PRIOR APPLICATION NUMBER: 08/710802
PRIOR FILING DATE: 1996-09-23
PRIOR APPLICATION NUMBER: 08/800699
PRIOR FILING DATE: 1997-02-14
PRIOR APPLICATION NUMBER: 08/828683
PRIOR FILING DATE: 1997-03-31
PRIOR APPLICATION NUMBER: 08/829270
PRIOR FILING DATE: 1997-03-31
PRIOR APPLICATION NUMBER: 08/928069
PRIOR FILING DATE: 1997-09-11
PRIOR APPLICATION NUMBER: 08/934494
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 09/143068
PRIOR FILING DATE: 1998-08-28
PRIOR APPLICATION NUMBER: 09/143707
PRIOR FILING DATE: 1998-08-28
PRIOR APPLICATION NUMBER: 09/151889
PRIOR FILING DATE: 1998-09-11
PRIOR APPLICATION NUMBER: 09/169104
PRIOR FILING DATE: 1998-10-09
PRIOR APPLICATION NUMBER: 09/202089
PRIOR FILING DATE: 1998-12-08
PRIOR APPLICATION NUMBER: 09/254311
PRIOR FILING DATE: 1999-03-03
PRIOR APPLICATION NUMBER: 09/304003
PRIOR FILING DATE: 1999-04-30
PRIOR APPLICATION NUMBER: 09/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380138
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380139
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/403297
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/423844
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 09/511133
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 09/511631
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 09/664610
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: 09/665350
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: 09/690169
PRIOR FILING DATE: 2000-10-16
PRIOR APPLICATION NUMBER: 09/690189
PRIOR FILING DATE: 2000-10-16
PRIOR APPLICATION NUMBER: 09/709238
PRIOR FILING DATE: 2000-11-18
PRIOR APPLICATION NUMBER: 09/866034
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 09/872035
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 09/884733
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: 09/886342
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: 09/866028
PRIOR FILING DATE: 2001-08-25
PRIOR APPLICATION NUMBER: PCT/US97/05230
PRIOR FILING DATE: 1997-03-31
PRIOR APPLICATION NUMBER: PCT/US98/19094
PRIOR FILING DATE: 1998-09-14

```

```

PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: PCT/US98/21407
PRIOR FILING DATE: 1998-10-09
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: 1998-12-01
PRIOR APPLICATION NUMBER: PCT/US99/05028
PRIOR FILING DATE: 1999-03-08
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: PCT/US99/20111
PRIOR FILING DATE: 1999-09-01
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28301
PRIOR FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: PCT/US99/28634
PRIOR FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: PCT/US00/04341
PRIOR FILING DATE: 2000-02-18
PRIOR APPLICATION NUMBER: PCT/US00/04342
PRIOR FILING DATE: 2000-02-18
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/06884
PRIOR FILING DATE: 2000-03-15
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: PCT/US00/13705
PRIOR FILING DATE: 2000-05-17
PRIOR APPLICATION NUMBER: PCT/US00/14941
PRIOR FILING DATE: 2000-05-30
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/20710
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: PCT/US01/17800
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: PCT/US01/19692
PRIOR FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: PCT/US01/21066
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: PCT/US01/21735
PRIOR FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 258
SEQ ID NO 44
LENGTH: 263
Query Match 99.8%; Score 1385; DB 23; Length 263;
Best Local Similarity 99.6%; Pred. No. 3.6e-128;
Matches 261; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
1 LLAWAQAFIVSNMLLAAYVSGGCFWMDNGHLVREDQTSAPGIRCLNWLDAOSGLASAPV 60
2 LLAWAQAFIVSNMLLAAYVSGGCFWMDNGHLVREDQTSAPGIRCLNWLDAOSGLASAPV 61
61 SGAGNHSYCNPNDEDPGPGWCYVSGEAGVPEKRPCEDLRCPEPTTSQALPAFTTEIOEASE 120
62 SGAGNHSYCNPNDEDPGPGWCYVSGEAGVPEKRPCEDLRCPEPTTSQALPAFTTEIOEASE 121
121 GPGADVQVFAFANALPARSEAAVQPVIGISQVRMNSKEKKDGLGTGLGVITMVI 180
```

```

|||||
DB 122 GPGADVQVFAFANALPARSEAAVQPVIGISQVRMNSKEKKDGLGTGLGVITMVI 181
QY 181 IAIAGAIILIGSYKRGKDLKECHDKVCEREMORITLPLSAFTPTCEIYDEKTVVHTS 240
DB 182 IAIAGAIILIGSYKRGKDLKECHDKVCEREMORITLPLSAFTPTCEIYDEKTVVHTS 241
QY 241 QTPVDPQSGSTPLMGQAGTPGA 262
DB 242 QTPVDPQSGSTPLMGQAGTPGA 263
```

```

RESULT 11
US-10-210-951-44
Sequence 44, Application US/10210951
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Masters, Scot A.
APPLICANT: Pan, James
APPLICANT: Pitti, Robert M.
APPLICANT: Roy, Margaret Ann
APPLICANT: Smith, Victoria
APPLICANT: Stone, Donna M.
APPLICANT: Watanabe, Colin K.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
FILE REFERENCE: P2931R1C1
CURRENT APPLICATION NUMBER: US/10/210,951
CURRENT FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: 60/014699
PRIOR FILING DATE: 1996-04-01
PRIOR APPLICATION NUMBER: 60/026943
PRIOR FILING DATE: 1996-09-23
PRIOR APPLICATION NUMBER: 60/059121
PRIOR FILING DATE: 1997-07-17
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/062037
PRIOR FILING DATE: 1997-10-10
PRIOR APPLICATION NUMBER: 60/063755
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063045
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063046
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/066511
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066772
PRIOR FILING DATE: 1997-11-24
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 258
SEQ ID NO 44
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
US-10-210-951-44
```

```

Query Match 99.8%; Score 1385; DB 26; Length 263;
Best Local Similarity 99.6%; Pred. No. 3.6e-128;
Matches 261; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```

1 LLAWAQAFIVSNMLLAAYVSGGCFWMDNGHLVREDQTSAPGIRCLNWLDAOSGLASAPV 60
2 LLAWAQAFIVSNMLLAAYVSGGCFWMDNGHLVREDQTSAPGIRCLNWLDAOSGLASAPV 61
61 SGAGNHSYCNPNDEDPGPGWCYVSGEAGVPEKRPCEDLRCPEPTTSQALPAFTTEIOEASE 120
62 SGAGNHSYCNPNDEDPGPGWCYVSGEAGVPEKRPCEDLRCPEPTTSQALPAFTTEIOEASE 121
```

QY 121 GPGADEVOVFAPANALPARSEAAVQPIGISOVRNMSKEKDLGTLGYLGTMMVII 180
DB 122 GPGADEVOVFAPANALPARSEAAVQPIGISOVRNMSKEKDLGTLGYLGTMMVII 181
QY 181 IAIAGIIIGSYRGKDLKEQHDQKVCEREMORITLPLSAFTNPTCEIYDEKTVVHTS 240
DB 182 IAIAGIIIGSYRGKDLKEQHDQKVCEREMORITLPLSAFTNPTCEIYDEKTVVHTS 241
QY 241 QTPVDPQEGSTPLMGQAGTPGA 262
DB 242 QTPVDPQEGSTPLMGQAGTPGA 263

RESULT 12

US-10-211-858-44
; Sequence 44, Application US/10211858
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Marsters, Scot A.
; APPLICANT: Pan, James
; APPLICANT: Pitti, Robert M.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stone, Donna M.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
; FILE REFERENCE: P2931R1C1
; CURRENT APPLICATION NUMBER: US/10/211,858
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/014699
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 60/026943
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 60/059121
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/062037
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 258
; SEQ ID NO 44
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-858-44

Query Match 99.8%; Score 1385; DB 26; Length 263;
Best Local Similarity 99.6%; Pred. No. 3, 6e-128;
Matches 261; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 LLAWVOAFLVSNMLLAAYVGGGCFWNGHLYREDQTSAPAGRLCINWMLDAOSGLASAPV 60
DB 2 LLAWVOAFLVSNMLLAAYVGGGCFWNGHLYREDQTSAPAGRLCINWMLDAOSGLASAPV 61
QY 61 SGAGNHSYCRNPEDPRGPMCVYSGAGVPEKPCEDLRCPETTSQALPAFTTIEOASE 120
DB 62 SGAGNHSYCRNPEDPRGPMCVYSGAGVPEKPCEDLRCPETTSQALPAFTTIEOASE 121

QY 121 GPGADEVOVFAPANALPARSEAAVQPIGISOVRNMSKEKDLGTLGYLGTMMVII 180
DB 122 GPGADEVOVFAPANALPARSEAAVQPIGISOVRNMSKEKDLGTLGYLGTMMVII 181
QY 181 IAIAGIIIGSYRGKDLKEQHDQKVCEREMORITLPLSAFTNPTCEIYDEKTVVHTS 240
DB 182 IAIAGIIIGSYRGKDLKEQHDQKVCEREMORITLPLSAFTNPTCEIYDEKTVVHTS 241
QY 241 QTPVDPQEGSTPLMGQAGTPGA 262
DB 242 QTPVDPQEGSTPLMGQAGTPGA 263

RESULT 13

US-10-211-884-44
; Sequence 44, Application US/10211884
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Marsters, Scot A.
; APPLICANT: Pan, James
; APPLICANT: Pitti, Robert M.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stone, Donna M.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
; FILE REFERENCE: P2931R1C1
; CURRENT APPLICATION NUMBER: US/10/211,884
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/014699
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 60/026943
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 60/059121
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/062037
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 258
; SEQ ID NO 44
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-884-44

Query Match 99.8%; Score 1385; DB 26; Length 263;
Best Local Similarity 99.6%; Pred. No. 3, 6e-128;
Matches 261; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 LLAWVOAFLVSNMLLAAYVGGGCFWNGHLYREDQTSAPAGRLCINWMLDAOSGLASAPV 60
DB 2 LLAWVOAFLVSNMLLAAYVGGGCFWNGHLYREDQTSAPAGRLCINWMLDAOSGLASAPV 61
QY 61 SGAGNHSYCRNPEDPRGPMCVYSGAGVPEKPCEDLRCPETTSQALPAFTTIEOASE 120

```

Db 62 SGAGNHSYCRNDEDEPRGWCYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASE 121
Qy 121 GGADEVQVFAPANALPARSEAAAVQPVIGISQVRMNSKEKKDLGTLGYVLGITMWYII 180
Db 122 GGADEVQVFAPANALPARSEAAAVQPVIGISQVRMNSKEKKDLGTLGYVLGITMWYII 181
Qy 181 IAIAGGIIIGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIIVDEKTVVHTS 240
Db 182 IAIAGGIIIGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIIVDEKTVVHTS 241
Qy 241 QTPVDPQEGSTPLMGQAGTPGA 262
Db 242 QTPVDPQEGSTPLMGQAGTPGA 263

RESULT 14
US-60-230-435-1580
; Sequence 1580, Application US/60230435
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: C1000768
; CURRENT APPLICATION NUMBER: US/60/230,435
; NUMBER OF SEQ ID NOS: 2991
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1580
; LENGTH: 263
; TYPE: PRP
; ORGANISM: HUMAN
US-60-230-435-1580

Query Match 99.8%; Score 1385; DB 27; Length 263;
Best Local Similarity 99.6%; Pred. No. 3.6e-128;
Matches 261; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLAWQAFIVSNMLLAAYSGGCFWMDNGHLVREDQTSAPAGLRCLNMLDAOSGLASAPV 60
Db 2 LLAWQAFIVSNMLLAAYSGGCFWMDNGHLVREDQTSAPAGLRCLNMLDAOSGLASAPV 61
Qy 61 SGAGNHSYCRNDEDEPRGWCYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASE 120
Db 62 SGAGNHSYCRNDEDEPRGWCYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASE 121
Qy 121 GGADEVQVFAPANALPARSEAAAVQPVIGISQVRMNSKEKKDLGTLGYVLGITMWYII 180
Db 122 GGADEVQVFAPANALPARSEAAAVQPVIGISQVRMNSKEKKDLGTLGYVLGITMWYII 181
Qy 181 IAIAGGIIIGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIIVDEKTVVHTS 240
Db 182 IAIAGGIIIGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIIVDEKTVVHTS 241
Qy 241 QTPVDPQEGSTPLMGQAGTPGA 262
Db 242 QTPVDPQEGSTPLMGQAGTPGA 263

RESULT 15
US-60-389-987-244
; Sequence 244, Application US/60389987
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465P2

```

```

; CURRENT APPLICATION NUMBER: US/60/389,987
; CURRENT FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 3025
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 244
; LENGTH: 263
; TYPE: PRP
; ORGANISM: Homo sapiens
US-60-389-987-244

Query Match 99.8%; Score 1385; DB 27; Length 263;
Best Local Similarity 99.6%; Pred. No. 3.6e-128;
Matches 261; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLAWQAFIVSNMLLAAYSGGCFWMDNGHLVREDQTSAPAGLRCLNMLDAOSGLASAPV 60
Db 2 LLAWQAFIVSNMLLAAYSGGCFWMDNGHLVREDQTSAPAGLRCLNMLDAOSGLASAPV 61
Qy 61 SGAGNHSYCRNDEDEPRGWCYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASE 120
Db 62 SGAGNHSYCRNDEDEPRGWCYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASE 121
Qy 121 GGADEVQVFAPANALPARSEAAAVQPVIGISQVRMNSKEKKDLGTLGYVLGITMWYII 180
Db 122 GGADEVQVFAPANALPARSEAAAVQPVIGISQVRMNSKEKKDLGTLGYVLGITMWYII 181
Qy 181 IAIAGGIIIGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIIVDEKTVVHTS 240
Db 182 IAIAGGIIIGYSYKRGKDLKEQHDQKVCEREMORITLPLSAFTNPCEIIVDEKTVVHTS 241
Qy 241 QTPVDPQEGSTPLMGQAGTPGA 262
Db 242 QTPVDPQEGSTPLMGQAGTPGA 263

```

Search completed: April 7, 2003, 08:49:48
Job time: 328 secs

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 08:42:52 ; Search time 31 Seconds
(without alignments)

1278.158 Million cell updates/sec

Title: US-10-057-951-2_COPY_2_263

Perfect score: 1388
Sequence: 1 LLAWVQAFVLSNMLAEAYG.....PVDPEGSTPLMGQAGTPGA 262

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 679521 seqs, 151232488 residues

Total number of hits satisfying chosen parameters: 679521

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Pending Patents_AA_New:*
1: /cgn2_6/prodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/prodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/prodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/prodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep:*
7: /cgn2_6/prodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1379	99.4	263	6 US-10-218-140-6002	Sequence 6002, Ap
2	1115.5	80.4	264	6 US-10-144-779-488	Sequence 488, Ap
3	154.5	11.1	293	7 US-60-452-680-14406	Sequence 14406, A
4	154.5	11.1	293	7 US-60-453-135-8958	Sequence 8958, Ap
5	154.5	11.1	293	7 US-60-453-050-8958	Sequence 8958, Ap
6	154.5	11.1	516	7 US-60-452-680-14405	Sequence 14405, A
7	154.5	11.1	516	7 US-60-453-135-8957	Sequence 8957, Ap
8	154.5	11.1	516	7 US-60-453-050-8957	Sequence 8957, Ap
9	154.5	11.1	527	5 US-09-612-314A-51	Sequence 51, Appl
10	154.5	11.1	562	1 PCT-US02-32263-26	Sequence 26, Appl
11	154.5	11.1	562	5 US-09-703-695A-4	Sequence 4, Appl
12	154.5	11.1	562	6 US-10-287-994-26	Sequence 26, Appl
13	154.5	11.1	562	7 US-60-452-680-14408	Sequence 14408, A
14	154.5	11.1	562	7 US-60-453-135-8960	Sequence 8960, Ap
15	154.5	11.1	562	7 US-60-453-050-8960	Sequence 8960, Ap
16	154.5	11.1	587	5 US-09-949-016-11501	Sequence 11501, A
17	146.5	10.6	655	1 PCT-US02-19017-28	Sequence 28, Appl
18	146.5	10.6	656	7 US-60-453-135-11516	Sequence 11516, A
19	146.5	10.6	656	7 US-60-453-050-11516	Sequence 11516, A
20	146	10.5	322	1 PCT-US02-27855-21	Sequence 21, Appl
21	146	10.5	322	1 PCT-US02-27855A-21	Sequence 21, Appl
22	146	10.5	322	6 US-10-233-675A-21	Sequence 21, Appl
23	143.5	10.3	482	7 US-60-452-680-14407	Sequence 14407, A
24	143.5	10.3	482	7 US-60-453-135-8959	Sequence 8959, Ap
25	143.5	10.3	482	7 US-60-453-050-8959	Sequence 8959, Ap
26	143	10.3	713	5 US-09-949-016-9983	Sequence 9983, Ap

27	143	10.3	713	7 US-60-452-680-22399	Sequence 22399, A
28	143	10.3	713	7 US-60-453-135-13847	Sequence 13847, A
29	143	10.3	713	7 US-60-453-050-13847	Sequence 13847, A
30	142	10.2	339	6 US-10-401-108-3	Sequence 3, Appl
31	142	10.2	363	6 US-10-292-418-11	Sequence 11, Appl
32	142	10.2	391	1 PCT-US02-37879-7	Sequence 7, Appl
33	142	10.2	391	6 US-10-304-287-7	Sequence 7, Appl
34	142	10.2	394	1 PCT-US02-37879-8	Sequence 8, Appl
35	142	10.2	394	6 US-10-304-287-8	Sequence 8, Appl
36	142	10.2	453	5 US-09-291-200A-1	Sequence 1, Appl
37	142	10.2	458	5 US-09-946-893B-4	Sequence 4, Appl
38	142	10.2	569	5 US-09-946-893B-5	Sequence 5, Appl
39	142	10.2	571	5 US-09-946-893B-8	Sequence 8, Appl
40	142	10.2	576	5 US-09-946-893B-6	Sequence 6, Appl
41	142	10.2	791	1 PCT-US02-37879-1	Sequence 1, Appl
42	142	10.2	791	6 US-10-304-287-1	Sequence 1, Appl
43	142	10.2	810	5 US-09-291-200A-5	Sequence 5, Appl
44	142	10.2	810	5 US-09-946-893B-2	Sequence 2, Appl
45	141.5	10.2	688	1 PCT-US02-27855-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-10-218-140-6002
; Sequence 6002, Application US/10218140
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES
; FILE REFERENCE: 15966-543 CON
; CURRENT APPLICATION NUMBER: US/10/218, 140
; PRIOR FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 09/540, 763
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/127, 728
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/127, 636
; PRIOR FILING DATE: 1999-04-02
; PRIOR APPLICATION NUMBER: 60/127, 607
; NUMBER OF SEQ ID NOS: 6322
; SOFTWARE: Curanator Version 1.0
; SEQ ID NO 6002
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-218-140-6002

Query Match 99.4%; Score 1379; DB 6; Length 263;
Best Local Similarity 99.2%; Pred. No. 1.5e-110;
Matches 260; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY	1	LLAWVQAFVLSNMLAEAYGSGGCFMDNGLHYREDOTSAPARCLMWDLOSGLASAPV	60
DB	2	LLAWVQAFVLSNMLAEAYGSGGCFMDNGLHYREDOTSAPARCLMWDLOSGLASAPV	61
QY	61	SGAGNHSYCNPEDEDRGFCWYSGAGVBEKPCEDIRCPETTSQALPAFTTEIOEASE	120
DB	62	SGAGNHSYCNPEDEDRGFCWYSGAGVBEKPCEDIRCPETTSQALPAFTTEIOEASE	121
QY	121	GGAGDEVQVAPANALPARSEAAVOPVIGISORVNMSSKKDGLGTIGYVIGIMVYI	180
DB	122	GGAGDEVQVAPANALPARSEAAVOPVIGISORVNMSSKKDGLGTIGYVIGIMVYI	181
QY	181	IAIGAGIILGYSYKSKDILKEOHDKVCEKEMQRTPLPSAFTNPTCEIVDEKTVVHTS	240
DB	182	IAIGAGIILGYSYKSKDILKEOHDKVCEKEMQRTPLPSAFTNPTCEIVDEKTVVHTS	241
QY	241	QTPVDPEGSTPLMGQAGTPGA 262	

Db 242 QTPVDPQEGTTPIMGQAGTPGA 263

RESULT 2

US-10-144-779-488
; Sequence 488, Application US/10144779
; GENERAL INFORMATION:
; APPLICANT: SUBRAMANIAN, Mani et al.
; TITLE OF INVENTION: MOUSE ORTHOLOGS OF HUMAN DISEASE GENES,
; TITLE OF INVENTION: PROTEINS ENCODED BY THESE MOUSE GENES, AND USES THEREOF
; FILE REFERENCE: C0001235
; CURRENT APPLICATION NUMBER: US/10/144, 779
; CURRENT FILING DATE: 2002-05-15
; NUMBER OF SEQ ID NOS: 864
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 488
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-144-779-488

Query Match 80.4%; Score 115.5; DB 6; Length 264;
Best Local Similarity 79.1%; Pred. No. 67e-88;
Matches 208; Conservative 17; Mismatches 37; Indels 1; Gaps 1;

Qy 1 LIATVQALVSNMILAEAYSGGCFWMDNGHLVREDQTSAPAGLRCLNWLDAOSGLASAPV 60

Db 2 LIATVHTFLSMLAEAYSGGCFWMDNGHLVREDQTSAPAGLRCLNWLDAOSGLASAPV 61

Qy 61 SGAGNHSYCRNPDEPRGPMCVSGEAGVPEKRCEDLRCPETTSQ-LPAFTTEIOEAS 119

Db 62 PBRGMHNTCRNPDDPRGPMCVISSETGVPEKRCEDVSCPETTSQAPPPSSAMELEBKS 121

Qy 120 EGPGEDEVQVPAANALPARSEAAVQPVIGISQVVMNSKEKKDGLTGLVIGITMVT 179

Db 122 GAPGKEAQVFPFANALPARSEAAVQPVIGISQVVMNSKEKKDGLTGLVIGITMVT 181

Qy 180 IIAIGAGIILGYSYRGDKLEKQKVCEREMQRTLLPLSAFNPTCEIVDEKTVVHT 239

Db 182 ILAIGAGIIVGTYTGRGDKLEKQKVCEREMQRTLLPLSAFNPTCEIVDEKTVVHT 241

Qy 240 SQTVPDQEGSTPLMGQAGTPGA 262

Db 242 NQTPADVOEGSTLLTGQAGTPGA 264

RESULT 3

US-60-452-680-14406
; Sequence 14406, Application US/60452680
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: GRUPE, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001450
; CURRENT APPLICATION NUMBER: US/60/452,680
; CURRENT FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14406
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-452-680-14406

Query Match 11.1%; Score 154.5; DB 7; Length 293;
Best Local Similarity 39.6%; Pred. No. 3e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

Qy 24 CFMDNGHLVREDQTSAPAGLRCLNWLDAOSGLASAPV-----GAGNHSYCRNPDE 74

Db 127 CYEDQGISYRGTWSTAESGAECTNM--NSSALAQKPYSGRRPDAIRLGLGNHNYCRNPDR 184

Qy 75 DPRGPMCVSGEAGVPEKRCEDLRCPETTS 105

Db 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 213

RESULT 4

US-60-453-135-8958
; Sequence 8958, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8958
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-8958

Query Match 11.1%; Score 154.5; DB 7; Length 293;
Best Local Similarity 39.6%; Pred. No. 3e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

Qy 24 CFMDNGHLVREDQTSAPAGLRCLNWLDAOSGLASAPV-----GAGNHSYCRNPDE 74

Db 127 CYEDQGISYRGTWSTAESGAECTNM--NSSALAQKYSORRRPDAIRLGLGNHNYCRNPDR 184

Qy 75 DPRGPMCVSGEAGVPEKRCEDLRCPETTS 105

Db 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 213

RESULT 5

US-60-453-050-8958
; Sequence 8958, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8958
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-8958

Query Match 11.1%; Score 154.5; DB 7; Length 293;
Best Local Similarity 39.6%; Pred. No. 3e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

Qy 24 CFMDNGHLVREDQTSAPAGLRCLNWLDAOSGLASAPV-----GAGNHSYCRNPDE 74

Db 127 CYEDQGISYRGTWSTAESGAECTNM--NSSALAQKPYSGRRPDAIRLGLGNHNYCRNPDR 184

Qy 75 DPRGPMCVSGEAGVPEKRCEDLRCPETTS 105

Db 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 213

RESULT 6

US-60-452-680-14405
; Sequence 14405, Application US/60452680
; GENERAL INFORMATION:

```
; APPLICANT: CARGILL, Michele
; APPLICANT: GRUPE, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001450
; CURRENT APPLICATION NUMBER: US/60/452,680
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 14405
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-452-680-14405

Query Match          11.1%; Score 154.5; DB 7; Length 516;
Best Local Similarity 39.6%; Pred. No. 5.9e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 24 CFWDNGHLYREDQTSAPAGRLCLNWLDAQSLASAPVS-----GAGNHSYCRNPDE 74
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 81 CYEDQGISYRGWTSTASGACTNW--NSSALAQKPYSGRRPDALRLGLGNHNYCHNPDR 138

QY 75 DPRGPMCVSAGEAGVPEKRPCEDLRCPEPTTS 105
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 139 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 167

RESULT 7
US-60-453-135-8957
; Sequence 8957, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8957
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-8957

Query Match          11.1%; Score 154.5; DB 7; Length 516;
Best Local Similarity 39.6%; Pred. No. 5.9e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 24 CFWDNGHLYREDQTSAPAGRLCLNWLDAQSLASAPVS-----GAGNHSYCRNPDE 74
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 81 CYEDQGISYRGWTSTASGACTNW--NSSALAQKPYSGRRPDALRLGLGNHNYCHNPDR 138

QY 75 DPRGPMCVSAGEAGVPEKRPCEDLRCPEPTTS 105
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 139 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 167

RESULT 8
US-60-453-050-8957
; Sequence 8957, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FASTSEQ for Windows Version 4.0
```

```
; SEQ ID NO 8957
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-8957

Query Match          11.1%; Score 154.5; DB 7; Length 516;
Best Local Similarity 39.6%; Pred. No. 5.9e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 24 CFWDNGHLYREDQTSAPAGRLCLNWLDAQSLASAPVS-----GAGNHSYCRNPDE 74
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 81 CYEDQGISYRGWTSTASGACTNW--NSSALAQKPYSGRRPDALRLGLGNHNYCHNPDR 138

QY 75 DPRGPMCVSAGEAGVPEKRPCEDLRCPEPTTS 105
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 139 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 167

RESULT 9
US-09-612-314A-51
; Sequence 51, Application US/09612314A
; GENERAL INFORMATION:
; APPLICANT: SMITH, RICHARD ANTHONY GODWIN
; APPLICANT: DODD, IAN
; APPLICANT: MOSSAKOWSKA, DANUTA EWA IRENA
; TITLE OF INVENTION: CONJUGATES OF SOLUBLE PEPTIDIC COMPOUNDS WITH
; TITLE OF INVENTION: MEMBRANE-BINDING AGENTS
; FILE REFERENCE: 37945-0004
; CURRENT APPLICATION NUMBER: US/09/612,314A
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 09/214,913
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: PCT/EP97/03715
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: GB 96 148 71.3
; PRIOR FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: tissue plasminogen activator
US-09-612-314A-51

Query Match          11.1%; Score 154.5; DB 5; Length 527;
Best Local Similarity 39.6%; Pred. No. 6e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 24 CFWDNGHLYREDQTSAPAGRLCLNWLDAQSLASAPVS-----GAGNHSYCRNPDE 74
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 92 CYEDQGISYRGWTSTASGACTNW--NSSALAQKPYSGRRPDALRLGLGNHNYCHNPDR 149

QY 75 DPRGPMCVSAGEAGVPEKRPCEDLRCPEPTTS 105
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 150 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 178

RESULT 10
PCT-US02-32263-26
; Sequence 26, Application PC/TUS0232263
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: DePrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Bower, Caryn
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; TITLE OF INVENTION: REMODELING AND GLYCOCONJUGATION OF PEPTIDES
; FILE REFERENCE: 040653-01-5050MO
```

```

; CURRENT APPLICATION NUMBER: PCT/US02/32263
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/338,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/334,233
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/334,301
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 26
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-32263-26

Query Match
Best Local Similarity 11.1%; Score 154.5; DB 1; Length 562;
Best Local Similarity 39.6%; Pred. No. 6.5e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 24 CFWDNGHLYREDQTSRPAAGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 74
DB 127 CYEDGISTRGWSTABSGAECTNM--NSSALAKPYSGRRPDALRLGNNHYCRNPDR 184
DB 185 DSK-FWCYVF-KAGKYSSEFCSTPACSEGNS 213

RESULT 11
US-09-703-695A-4
; Sequence 4, Application US/09703695A
; GENERAL INFORMATION:
; APPLICANT: Xu, Yuan
; TITLE OF INVENTION: REVERSE-PHASE HPLC ASSAY FOR PLASMINOGEN ACTIVATORS
; FILE REFERENCE: P1789R1
; CURRENT APPLICATION NUMBER: US/09/703,695A
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: US 60/163,607
; PRIOR FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 4
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-703-695A-4

Query Match
Best Local Similarity 11.1%; Score 154.5; DB 5; Length 562;
Best Local Similarity 39.6%; Pred. No. 6.5e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 24 CFWDNGHLYREDQTSRPAAGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 74
DB 127 CYEDGISTRGWSTABSGAECTNM--NSSALAKPYSGRRPDALRLGNNHYCRNPDR 184
QY 75 DRGPMCYVSGEAGVPEKRPCEDLRCPEPTTS 105
DB 185 DSK-FWCYVF-KAGKYSSEFCSTPACSEGNS 213

RESULT 12
```

```

US-10-287-994-26
; Sequence 26, Application US/10287994
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: Neese, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Bower, Caryn
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; TITLE OF INVENTION: REMODELING AND GLYCOCONJUGATION OF PEPTIDES
; FILE REFERENCE: 040853-01-5052-00
; CURRENT APPLICATION NUMBER: US/10/287,994
; CURRENT FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 26
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-287-994-26

Query Match
Best Local Similarity 11.1%; Score 154.5; DB 6; Length 562;
Best Local Similarity 39.6%; Pred. No. 6.5e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 24 CFWDNGHLYREDQTSRPAAGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 74
DB 127 CYEDGISTRGWSTABSGAECTNM--NSSALAKPYSGRRPDALRLGNNHYCRNPDR 184
QY 75 DRGPMCYVSGEAGVPEKRPCEDLRCPEPTTS 105
DB 185 DSK-FWCYVF-KAGKYSSEFCSTPACSEGNS 213

RESULT 13
US-60-452-680-14408
; Sequence 14408, Application US/60452680
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: GRUBE, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CU001450
; CURRENT APPLICATION NUMBER: US/60/452,680
; CURRENT FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14408
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-452-680-14408

Query Match
Best Local Similarity 11.1%; Score 154.5; DB 7; Length 562;
Best Local Similarity 39.6%; Pred. No. 6.5e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 24 CFWDNGHLYREDQTSRPAAGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 74
```

Job time : 32 secs

```
Db 127 CFEODGISTYRGWTGTAESGAEACTNW--NSSALAQKPYSGRRPDATRLGIGNHNYCRNPDR 184
QY 75 DPRGWCYVSGEAGVPEKRPCEDLRCPETTS 105
Db 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 213
```

RESULT 14

```
US-60-453-135-8960
; Sequence 8960, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: C1001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8960
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-8960
```

Query Match 11.1%; Score 154.5; DB 7; Length 562;
Best Local Similarity 39.6%; Pred. No. 6.5e-05;

Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

```
QY 24 CFWDGHLTYREDQTPAPGLRCLNMLDAQSLASAPVS-----GAGNHSYCRNPDE 74
Db 127 CFEODGISTYRGWTGTAESGAEACTNW--NSSALAQKPYSGRRPDATRLGIGNHNYCRNPDR 184
QY 75 DPRGWCYVSGEAGVPEKRPCEDLRCPETTS 105
Db 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 213
```

RESULT 15

```
US-60-453-050-8960
; Sequence 8960, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: C1001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8960
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-8960
```

Query Match 11.1%; Score 154.5; DB 7; Length 562;
Best Local Similarity 39.6%; Pred. No. 6.5e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

```
QY 24 CFWDGHLTYREDQTPAPGLRCLNMLDAQSLASAPVS-----GAGNHSYCRNPDE 74
Db 127 CFEODGISTYRGWTGTAESGAEACTNW--NSSALAQKPYSGRRPDATRLGIGNHNYCRNPDR 184
QY 75 DPRGWCYVSGEAGVPEKRPCEDLRCPETTS 105
Db 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 213
```

Search completed: April 7, 2003, 08:50:27


```
US-08-811-949-63
; Sequence 63, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NINWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 472 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-811-949-63

Query Match      11.8%; Score 163.5; DB 2; Length 472;
Best Local Similarity 40.0%; Pred. No. 7.2e-09;
Matches 38; Conservative 6; Mismatches 38; Indels 13; Gaps 4;

QY 20 GSGGCFWMDNGHLVREDQTSAPGLRCLNWLDAQSLASAPVS-----GAGNHSYICR 70
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 33 GARSCEDQGISYRGWTWTASGABETNM--NSSALQKPYSGRRPDRIQLGNHNYICR 90
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 71 NPDEPRGFWCVSGEAGVPEKRPPCEDLRCPEPTTS 105
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 91 NPDRDSK-FWCYVF-KAGKYSSEFCSTPACSEGNS 123

RESULT 3
US-08-811-949-49
; Sequence 49, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NINWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESS: P. C.
```

```
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-811-949-49

Query Match      11.1%; Score 154.5; DB 2; Length 437;
Best Local Similarity 39.6%; Pred. No. 5.7e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 24 CFMDNGHLVREDQTSAPGLRCLNWLDAQSLASAPVS-----GAGNHSYICNPDE 74
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 2 CYEDQGISYRGWTWTASGABETNM--NSSALQKPYSGRRPDRIQLGNHNYICNPDR 59
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 75 DPRGFWCVSGEAGVPEKRPPCEDLRCPEPTTS 105
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 60 DSK-FWCYVF-KAGKYSSEFCSTPACSEGNS 88

RESULT 4
US-08-811-949-51
; Sequence 51, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NINWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESS: P. C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
```


NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-51

Query Match 11.1%; Score 154.5; DB 2; Length 437;
Best Local Similarity 39.6%; Pred. No. 5.7e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

OY 24 CFWDGHLVREDQTSPPAPGLRCLNMLDAQSGLASAPVS-----GAGNHSYCRNPDE 74
DB 2 CYEDQGISYRGTWSTAESGAECTNW--NSSALAQKPYSGRRPDARLGLGNHNYCRNPDR 59
OY 75 DRRGPMCVSGEAGVPEKRPCEDLRCPEPTTS 105
DB 60 DSK-PWCTVF-KAGKYSSEFCSTPACSEGN 88

RESULT 5
US-08-811-949-55
Sequence 55, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIMA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, YUJI
APPLICANT: KOBAYASHI, MASARAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-55

Query Match 11.1%; Score 154.5; DB 2; Length 437;

Best Local Similarity 39.6%; Pred. No. 5.7e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
OY 24 CFWDGHLVREDQTSPPAPGLRCLNMLDAQSGLASAPVS-----GAGNHSYCRNPDE 74
DB 2 CYEDQGISYRGTWSTAESGAECTNW--NSSALAQKPYSGRRPDARLGLGNHNYCRNPDR 59
OY 75 DRRGPMCVSGEAGVPEKRPCEDLRCPEPTTS 105
DB 60 DSK-PWCTVF-KAGKYSSEFCSTPACSEGN 88

RESULT 6
US-08-811-949-57
Sequence 57, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIMA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, YUJI
APPLICANT: KOBAYASHI, MASARAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-57

Query Match 11.1%; Score 154.5; DB 2; Length 437;
Best Local Similarity 39.6%; Pred. No. 5.7e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

OY 24 CFWDGHLVREDQTSPPAPGLRCLNMLDAQSGLASAPVS-----GAGNHSYCRNPDE 74
DB 2 CYEDQGISYRGTWSTAESGAECTNW--NSSALAQKPYSGRRPDARLGLGNHNYCRNPDR 59
OY 75 DRRGPMCVSGEAGVPEKRPCEDLRCPEPTTS 105
DB 60 DSK-PWCTVF-KAGKYSSEFCSTPACSEGN 88

RESULT 7
US-07-609-510B-16
Sequence 16, Application US/07609510B

```

; Patent No. 5326700
; GENERAL INFORMATION:
; APPLICANT: Berg et al.
; TITLE OF INVENTION: Method for Altering Post-Translational Processing of Tissue P
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN.
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB storage
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/609,510B
; FILING DATE: 19901106
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-609-510B-16

Query Match      11.1%; Score 154.5; DB 1; Length 527;
Best Local Similarity 39.6%; Pred. No. 7.6e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

Qy 24 CFWDNGHLREDDTSPAPGRCLNWLDAQSGLASAPVS-----GAGNHSYCRNPD 74
Db 92 CYEDGISTYRGTWSTAESGAECTNW--NSSALAKPYSGRRPDARLGLGNHNYCRNPD 149
Qy 75 DPGPMCVSGEAGVPEKRPCEDLRCPETTS 105
Db 150 DSK-PMCYVF-KAGKYSSEFCSTPACSEGNS 178

RESULT 8
US-08-811-949-39
; Sequence 39, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIMA, MINBO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESS: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-811-949-39

Query Match      11.1%; Score 154.5; DB 2; Length 527;
Best Local Similarity 39.6%; Pred. No. 7.6e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

Qy 24 CFWDNGHLREDDTSPAPGRCLNWLDAQSGLASAPVS-----GAGNHSYCRNPD 74
Db 92 CYEDGISTYRGTWSTAESGAECTNW--NSSALAKPYSGRRPDARLGLGNHNYCRNPD 149
Qy 75 DPGPMCVSGEAGVPEKRPCEDLRCPETTS 105
Db 150 DSK-PMCYVF-KAGKYSSEFCSTPACSEGNS 178

RESULT 9
PCT-US91-01025A-2
; Sequence 2, Application PC/TUS9101025A
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: Tissue Plasminogen Activator Having Fibrin
; TITLE OF INVENTION: Specific Properties
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/01025A
; FILING DATE: 19910214
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/486,657
; FILING DATE: 1 March 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Haasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 454P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/266-1896
; TELEFAX: 415/952-9881
; TELERX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; PCT-US91-01025A-2

Query Match      11.1%; Score 154.5; DB 5; Length 527;
Best Local Similarity 39.6%; Pred. No. 7.6e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
```

Qy 24 CFWDNGHLYREDQTSAPAGLCRLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 74
Db 92 CYEDQGISYRGTWSTAESGAECTNW--NSSALAQKPYSGRRPDAIRLGLGNHNYCRNPDR 149
Qy 75 DPRGPMCVYSGEAGVPEKRPCEDLRCPEPTTS 105
Db 150 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 178

RESULT 10
5185259-8
; Patent No. 5185259
; APPLICANT: GOEDEL, DAVID V.; KOHR, WILLIAM J.; PENNICA, DIANE;
; VEHAR, GORDON A.
; TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN
; ACTIVATOR
; NUMBER OF SEQUENCES: 15
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/489,855
; FILING DATE: 02-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 12,694
; FILING DATE: 09-FEB-1987
; APPLICATION NUMBER: 483,052
; FILING DATE: 07-APR-1983
; APPLICATION NUMBER: 398,003
; FILING DATE: 14-JUL-1982
; APPLICATION NUMBER: 374,860
; FILING DATE: 05-MAY-1982
; SEQ ID NO: 8
; LENGTH: 527
5185259-8
Query Match 11.1%; Score 154.5; DB 6; Length 527;
Best Local Similarity 39.6%; Pred. No. 7.6e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
Qy 24 CFWDNGHLYREDQTSAPAGLCRLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 74
Db 92 CYEDQGISYRGTWSTAESGAECTNW--NSSALAQKPYSGRRPDAIRLGLGNHNYCRNPDR 149
Qy 75 DPRGPMCVYSGEAGVPEKRPCEDLRCPEPTTS 105
Db 150 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 178

RESULT 11
5520913-1
; Patent No. 5520913
; APPLICANT: ANDERSON, STEPHEN; BENNETT, WILLIAM F.; BORTSTEIN,
; DAVID; HIGGINS, DEBORAH L.; PRONI, NICHOLAS F.; ZOLLER, MARK J.
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR HAVING
; ZYMOGENIC PROPERTIES
; NUMBER OF SEQUENCES: 35
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/88,451
; FILING DATE: 06-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 770,510
; FILING DATE: 03-OCT-1991
; APPLICATION NUMBER: 384,608
; FILING DATE: 24-JUL-1989
; APPLICATION NUMBER: 240,856
; FILING DATE: 02-SEP-1988
; SEQ ID NO: 1
; LENGTH: 527
5520913-1

Query Match 11.1%; Score 154.5; DB 6; Length 527;
Best Local Similarity 39.6%; Pred. No. 7.6e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

Qy 24 CFWDNGHLYREDQTSAPAGLCRLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 74
Db 92 CYEDQGISYRGTWSTAESGAECTNW--NSSALAQKPYSGRRPDAIRLGLGNHNYCRNPDR 149
Qy 75 DPRGPMCVYSGEAGVPEKRPCEDLRCPEPTTS 105
Db 150 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 178

RESULT 12
5200340-6
; Patent No. 5200340
; APPLICANT: FOSTER, DONALD C.; MUIVITHILL, EILEEN R.; O'HARA,
; PATRICK J.; PINDEL, KURT; YOSHITAKE, SHINJI
; TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
; ACTIVATORS
; NUMBER OF SEQUENCES: 34
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/53,412
; FILING DATE: 22-MAY-1987
; SEQ ID NO: 6
; LENGTH: 546
5200340-6
Query Match 11.1%; Score 154.5; DB 6; Length 546;
Best Local Similarity 39.6%; Pred. No. 8e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
Qy 24 CFWDNGHLYREDQTSAPAGLCRLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 74
Db 127 CYEDQGISYRGTWSTAESGAECTNW--NSSALAQKPYSGRRPDAIRLGLGNHNYCRNPDR 184
Qy 75 DPRGPMCVYSGEAGVPEKRPCEDLRCPEPTTS 105
Db 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 213

RESULT 13
US-08-811-949-43
; Sequence 43, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NITWA, MINO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000

TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 562 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-43

Query Match 11.1%; Score 154.5; DB 2; Length 562;
Best Local Similarity 39.6%; Pred. No. 8.4e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 24 CFWDNGHLTYREDQTPAPGLRCINWLDASGLASAPVS-----GAGNHSYCRNPDE 74
DB 127 CYEDQGISYRGTWSTASGAECTNW--NSSALAQPYSGRRPDAIRLGLGNHNYCRNPDR 184

QY 75 DPGPWCYVSGEAGVPEKRPCEDLRCPEPTTS 105
DB 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 213

RESULT 14
US-08-560-098A-50
Sequence 50, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WNEEDT, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins having fibrinolytic and
TITLE OF INVENTION: Coagulation-inhibiting Properties
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKewen, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 562 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-560-098A-50

Query Match 11.1%; Score 154.5; DB 2; Length 562;
Best Local Similarity 39.6%; Pred. No. 8.4e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
QY 24 CFWDNGHLTYREDQTPAPGLRCINWLDASGLASAPVS-----GAGNHSYCRNPDE 74
DB 127 CYEDQGISYRGTWSTASGAECTNW--NSSALAQPYSGRRPDAIRLGLGNHNYCRNPDR 184

DB 127 CYEDQGISYRGTWSTASGAECTNW--NSSALAQPYSGRRPDAIRLGLGNHNYCRNPDR 184
QY 75 DPGPWCYVSGEAGVPEKRPCEDLRCPEPTTS 105
DB 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 213

RESULT 15
US-08-883-795A-38
Sequence 38, Application US/08883795A
Patent No. 5985607
GENERAL INFORMATION:
APPLICANT: Delcove, Genevieve
APPLICANT: Awang, Gregor
TITLE OF INVENTION: Recombinant DNA Molecules and Expression
TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: BERESKIN & PARR
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3T2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/883,795A
FILING DATE: 27-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gravelle, Micheline
REGISTRATION NUMBER: 40,261
REFERENCE/DOCKET NUMBER: 7841-062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 562 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein (tpa)
US-08-883-795A-38

Query Match 11.1%; Score 154.5; DB 2; Length 562;
Best Local Similarity 39.6%; Pred. No. 8.4e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
QY 24 CFWDNGHLTYREDQTPAPGLRCINWLDASGLASAPVS-----GAGNHSYCRNPDE 74
DB 127 CYEDQGISYRGTWSTASGAECTNW--NSSALAQPYSGRRPDAIRLGLGNHNYCRNPDR 184
QY 75 DPGPWCYVSGEAGVPEKRPCEDLRCPEPTTS 105
DB 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 213

Search completed: April 7, 2003, 08:44:14
Job time : 17 secs

GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:25 / Search time 19.6216 Seconds
(without alignments)
1185.658 Million cell updates/sec

Title: US-10-057-951-2_COPY_22_263

Perfect score: 1289
Sequence: 1 SGCGFMDNGHLVREDQTSPLA.....PVDPOEGSTPLMGAQAGPGA 242

Scoring table: BLOSUM62
Gapop 10.0, Gapept 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	154.5	12.0	562	1 UKHUT	t-plasminogen acti
2	154	11.9	559	1 A35029	t-plasminogen acti
3	147.5	11.4	559	1 A29941	t-plasminogen acti
4	146.5	11.4	655	1 A46688	hepatocyte growth
5	145.5	11.3	291	2 I38098	t-plasminogen acti
6	142	11.0	810	1 PLHU	plasmin (EC 3.4.21
7	138	10.7	169	2 A40522	plasmin (EC 3.4.21
8	138	10.7	433	1 UKBAY	u-plasminogen acti
9	135.5	10.5	442	1 UKPG	u-plasminogen acti
10	134	10.4	431	2 US0599	t-plasminogen acti
11	134	10.4	477	1 A34369	t-plasminogen acti
12	134	10.4	477	2 US0598	t-plasminogen acti
13	133.5	10.4	716	1 J05061	macrophage-stimula
14	132	10.2	431	1 UKHU	u-plasminogen acti
15	130.5	10.1	716	1 A40332	macrophage-stimula
16	129	10.0	810	2 I46260	plasmin (EC 3.4.21
17	127	9.9	394	2 J06600	t-plasminogen acti
18	127	9.9	433	1 J00560	u-plasminogen acti
19	126.5	9.8	728	1 A60185	hepatocyte growth
20	124.5	9.7	810	2 B30848	plasmin (EC 3.4.21
21	124	9.6	622	1 TBHU	thrombin (EC 3.4.2
22	124	9.6	728	1 JH0579	hepatocyte growth
23	124	9.6	812	1 PLBO	plasmin (EC 3.4.21
24	123.5	9.6	434	1 A35005	u-plasminogen acti
25	123	9.5	593	2 S45281	coagulation factor
26	123.5	9.5	4548	1 S00657	apoptotrin(a) (EC
27	122	9.5	603	2 S28941	coagulation factor
28	121.5	9.4	728	1 A35644	hepatocyte growth
29	120	9.3	790	1 PLPG	plasmin (EC 3.4.21

30	119.5	9.3	477	2 J50597	t-plasminogen acti
31	119.5	9.3	560	1 JC4795	plasma hyaluronan
32	119	9.2	433	1 UKMS	u-plasminogen acti
33	118.5	9.2	615	1 KFHU12	coagulation factor
34	118	9.2	432	1 S18932	u-plasminogen acti
35	118	9.2	710	1 I51283	hepatocyte growth
36	117	9.1	625	1 TBBO	thrombin (EC 3.4.2
37	116	9.0	1420	2 A32869	apolipoprotein(a)
38	115.5	9.0	711	1 A47136	macrophage-stimula
39	114.5	8.9	812	1 PLMS	plasmin (EC 3.4.21
40	114	8.8	458	2 A35827	thrombin (EC 3.4.2
41	112	8.7	465	2 A61545	plasmin (EC 3.4.21
42	110.5	8.6	685	1 A46289	neutrotrophic recep
43	110	8.5	558	2 JC5878	plasma hyaluronan
44	108.5	8.4	123	2 C61545	plasmin (EC 3.4.21
45	107	8.3	617	2 S10511	thrombin (EC 3.4.2

ALIGNMENTS

RESULT 1

UKHUT
t-plasminogen activator (EC 3.4.21.68) precursor [validated] - human
N/Alternate names: t-PA; tissue plasminogen activator
C/Species: Homo sapiens (man)
C/Date: 14-Nov-1983 #sequence revision 14-Nov-1983 #text change 08-Dec-2000
C/Accession: A94004; A23529; J05062; A93293; S02125; A91343; A93951; A91322; A54645; I60
R/Ny, T.; Elgh, F.; Lund, B.
Proc. Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984
A/Title: The structure of the human tissue-type plasminogen activator gene: correlation
A/Reference number: A94004; MUID:84298137; PMID:6089198
A/Accession: A94004
A/Molecule type: DNA
A/Residues: 1-562 <NT>
A/Cross-references: GB:L00141
A/Note: The codon given for residue 93 (ACC) is inconsistent with the authors' translati.
R/Friener Degen, S.J.; Rajput, B.; Reich, E.
J. Biol. Chem. 261, 6972-6985, 1986
A/Title: The human tissue plasminogen activator gene.
A/Reference number: A23529; MUID:86196143; PMID:3009482
A/Accession: A23529
A/Molecule type: DNA
A/Residues: 1-562 <DE>
A/Cross-references: GB:K03021; NID:G339817; PIDN:AAA98809.1; PID:G339818
R/Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuda, S.; Higashio, K.
Agric. Biol. Chem. 55, 1225-1232, 1991
A/Title: Purification and characterization of tissue plasminogen activator secreted by h
A/Reference number: J05062; MUID:91291340; PMID:1368681
A/Accession: J05062
A/Molecule type: mRNA
A/Residues: 31-562 <IT>
A/Cross-references: DDBJ:D01096; NID:G220128; PIDN:BA00881.1; PID:G441174
A/Experimental source: embryonic lung fibroblast IMR-90 cells
A/Note: Part of this sequence, including the amino end of the mature protein, was confir
R/Pemica, D.; Holmes, W.E.; Kohr, W.J.; Harkins, R.N.; Vehar, G.A.; Ward, C.A.; Bennett
Nature 301, 214-221, 1983
A/Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Esche
A/Reference number: A93293; MUID:83115262; PMID:6337343
A/Accession: A93293
A/Molecule type: mRNA
A/Residues: 1-562 <PEN>
A/Cross-references: GB:L00141
A/Experimental source: melanoma cells
R/Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M.
Nucleic Acids Res. 16, 5695, 1988
A/Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human fe
A/Reference number: S02125; MUID:88262579; PMID:3133640
A/Accession: S02125
A/Status: translation not shown
A/Molecule type: mRNA
A/Residues: 1-562 <SAS>
A/Cross-references: EMBL:X07393; NID:G37243; PIDN:CMA0302.1; PID:G37244

A:Experimental source: fetal lung cells
 R:Kakitani, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Okada, K.; Ma
 ERBS Lett. 189, 145-149, 1985
 A:Title: Expression in Escherichia coli of finger-domain lacking tissue-type plasminogen
 A:Reference number: A91343; MUID:85285620; PMID:3896853
 A:Accession: A91343
 A:Molecule type: mRNA
 A:Residues: 1-38, 'G', 86-433, 'E', 435-562 <KAG>
 R:Edlund, T.; Ny, T.; Randy, M.; Hedén, L.O.; Palm, G.; Holmgren, E.; Josephson, S.
 Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983
 A:Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen acti
 A:Reference number: A93951; MUID:83169656; PMID:6572897
 A:Accession: A93951
 A:Molecule type: mRNA
 A:Residues: 251-358 <EDL>
 A:Experimental source: melanoma cells
 R:Polh, G.; Kallstrom, M.; Bergsdorf, N.; Wallen, P.; Jorvall, H.
 Biochemistry 23, 3701-3707, 1984
 A:Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived am
 differences.
 A:Reference number: A90488; MUID:85000468; PMID:6433976
 A:Contents: annotation; melanoma cells; partial sequence of residues 35-562, active and
 R:Polh, G.; Kaplan, L.; Einarsson, M.; Wallen, P.; Jorvall, H.
 FEBS Lett. 168, 29-32, 1984
 A:Title: Differences between uterine and melanoma forms of tissue plasminogen activator.
 A:Reference number: A91322; MUID:84158556; PMID:6538514
 A:Accession: A91322
 A:Molecule type: protein
 A:Residues: 33-45/311-320 <POH>
 A:Experimental source: uterus
 A:Note: in the uterus, cleavage of the activation peptide may also occur after 38-Gln
 R:van Zonneveld, A.J.; Veerman, H.; Pannekoek, H.
 J. Biol. Chem. 261, 14214-14218, 1986
 A:Reference number: A37567; MUID:87033611; PMID:3021732
 A:Contents: annotation; fibrin binding site
 R:Verheijen, J.H.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, P.H.; Enger
 EMBO J. 5, 3525-3530, 1986
 A:Title: Involvement of finger domain and kringles 2 domain of tissue-type plasminogen ac
 A:Reference number: A37568; MUID:87161761; PMID:3030730
 A:Contents: annotation; fibrin binding site
 R:Dodd, I.; Nunn, B.; Robinson, J.H.
 Thromb. Haemost. 59, 523-528, 1988
 A:Title: Isolation, identification and pharmacokinetic properties of human tissue-type P
 A:Reference number: A60902; MUID:89044681; PMID:3142086
 A:Contents: annotation; novel forms of expressed recombinant t-PA
 R:Harris, T.J.R.; Patel, T.; Marston, F.A.O.; Little, S.; Emtage, J.S.; Odenakker, G.;
 Mol. Biol. Med. 3, 279-292, 1986
 A:Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its exp
 A:Reference number: A54645; MUID:86284200; PMID:3090401
 A:Accession: A54645
 A:Molecule type: mRNA
 A:Residues: 1-562 <HAR>
 A:Cross-references: GB:M15518; NID:G190031; PIDN:AAA60111.1; PID:G190032
 A:Note: Parts of this sequence were confirmed by peptide sequencing
 R:Reddy, V.B.; Garramone, A.J.; Sasaak, H.; Wei, C.
 DNA 6, 461-472, 1987
 A:Title: Expression of human uterine tissue-type plasminogen activator in mouse cells us
 A:Reference number: I60110; MUID:88054470; PMID:2824147
 A:Accession: I60110
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-562 <RES>
 A:Cross-references: GB:M18182; NID:G340176; PIDN:AAA36800.1; PID:G340177
 R:Flaherty, R.; Waller, E.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schleuning, W.D.
 J. Biol. Chem. 260, 11223-11230, 1985
 A:Title: Isolation and characterization of the human tissue-type plasminogen activator B
 A:Reference number: I55232; MUID:85289338; PMID:3161893
 A:Accession: I55232
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-36 <RE2>
 A:Cross-references: GB:M11890; NID:G339837; PIDN:AAA61213.1; PID:G339839

C:Comment: Cleavage by plasmin or trypsin produces two chains held together by a single
 C:Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val bond. It
 C:Comment: t-PA binds chain A of fibrin by kringles 2 and the fibronectin type I repeat.
 C:Genetics:
 A:Gene: GDB:PLAT
 A:Cross-references: GDB:119496; OMIM:173370
 A:Map position: 8p12-8p12
 A:Introns: 24/3: 39/1: 85/1: 122/1: 180/2: 211/1: 268/2: 297/1: 362/2: 408/1: 455/3: 510
 C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
 C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringles; plasma; serine proteinase
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-32/Domain: signal sequence #status predicted <PRO>
 F:33-562/Product: propeptide #status predicted <PRO>
 F:33-310/Product: t-plasminogen activator #status experimental <MNT>
 F:41-78/Domain: fibronectin type I repeat homology <1F1>
 F:46-119/Domain: EGF homology <EGF>
 F:127-208/Domain: kringles homology <KR1>
 F:215-296/Domain: kringles homology <KR2>
 F:311-562/Product: t-plasminogen activator chain B #status experimental <BCH>
 F:41-71,68-78,86-97,91-108,110-119,127-208,148-190,179-203,215-226,236-278,267-291,299-4
 F:152,483/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:219/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
 F:310-311/Cleavage site: Arg-116 (plasmin, trypsin) #status experimental
 F:357,406/Active site: His, Asp #status predicted
 F:513/Active site: Ser #status experimental
 Query Match 12.0%; Score 154.5; DB 1; Length 562;
 Best Local Similarity 39.6%; Pred. No. 2.5e-05;
 Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
 Db 4 CPMDGHLVREDQTSFAPGRLCNWLDAGSGLASAPVS-----GAGNYSYCNPD 54
 127 CYEDQGISRTGTSTWESGAECTNM--NSSALAKRYSRRDRLGLGNHNYCRNPR 184
 QY 55 DPRGPMCVYSGEAVPEKRPCEDLRCPEETS 85
 Db 185 DSK-PMCVYF-KAGKYSSEKCPACSEGN 213
 RESULT 2
 A35029
 C-Plasminogen activator (BC 3.4.21.68) precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A35029; A31597
 R:Feng, P.; Ohlsson, M.; Ny, T.
 J. Biol. Chem. 265, 2022-2027, 1990
 A:Title: The structure of the TARA-less rat tissue-type plasminogen activator gene. Spec
 A:Reference number: A35029; MUID:90130448; PMID:2105315
 A:Accession: A35029
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-559 <FEN>
 A:Cross-references: GB:M1197; NID:G207429; PIDN:AAA42261.1; PID:G207431; GB:J05226
 R:Ny, T.; Leonardsson, G.; Heuvel, A.J.W.
 DNA 7, 671-677, 1988
 A:Title: Cloning and characterization of a cDNA for rat tissue-type plasminogen activato
 A:Reference number: A31597; MUID:89170114; PMID:3148445
 A:Accession: A31597
 A:Molecule type: mRNA
 A:Residues: 1-379, 'K', 381-559 <NYT>
 A:Cross-references: GB:M23697; NID:G530159; PIDN:AAA41812.1; PID:G530160
 C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
 C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringles; serine proteinase
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-29/Domain: propeptide #status predicted <PRO>
 F:30-559/Product: t-plasminogen activator #status predicted <MNT>
 F:30-108/Product: t-plasminogen activator chain A #status predicted <ACH>
 F:38-75/Domain: fibronectin type I repeat homology <1F1>
 F:83-116/Domain: EGF homology <EGF>
 F:124-205/Domain: kringles homology <KR1>
 F:213-294/Domain: kringles homology <KR2>

A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-291 <SIE>
A:Cross-references: EMBL:X13097; NID:g35282; PIDD:CAA31489.1; PID:g35283
C:Comment: For the main splice form, see PIR:UKHUT. This form probably does not have pro
C:Genetics:
A:Gene: GDB:PLAT
A:Cross-references: GDB:119496; OMIM:173370
A:Map position: 8p12-8p12
A:Intron: 24/3; 39/1; 65/1; 122/1; 180/2; 211/1; 268/2
C:Superfamily: tissue plasminogen activator, EGF homology, fibronectin type I repeat hom
C:Keywords: alternative splicing, fibrinolysis, glycoprotein, kringle
F:1-23/Domain: signal sequence #status predicted <SIG>
F:34-32/Domain: propeptide #status predicted <PRO>
F:33-39/Domain: t-plasminogen activator, inactive endothelial splice form #status pred
F:41-78/Domain: fibronectin type I repeat homology <EFG>
F:86-119/Domain: EGF homology <KRI>
F:127-208/Domain: kringle homology #status atypical <KR2>
F:215-291/Domain: kringle homology #status atypical <KR2>
F:41-71,69-78,86-97,91-108,110-119,127-208,148-190,179-203/Disulfide bonds: #status pred

Query Match 11.3%; Score 145.5; DB:2; Length 291;
Best Local Similarity 38.5%; Pred. No. 6;9e-05;
Matches 35; Conservative 6; Mismatches 37; Indels 13; Gaps 4;

Qy 4 CFWDNGHYRBDQTSPPAGRLCNWLDAGSGLASAPVS-----GAGNHSYCRNDE 54
Db 127 CYEDQGISYRGWSTAESGACTW--NSBALAQNAYSGRPDAIRLGLGNHNCYRPR 184
Qy 55 DRGPMCVYSGAGVPEKPECDLRCPENTTS 85
Db 185 DSK-FMCTVF-KAGKYSSEFCSPRACSEGN 213

RESULT 6
pLNU
plasmin (EC 3.4.21.7) precursor [validated] - human
N:Alternate names: plasminogen precursor [mismor]
N:Contains: angiotatin; microplasmin; plasminogen
C:Species: Homo sapiens (man)
C:Date: 24-Apr-1984 #sequence revision 02-Dec-1994 #text change 15-Sep-2000
C:Accession: A35229; I52242; A26646; 163738; 184609; S03375; A00922; A04627; A04625; A04
J:Biochem. 265, 6104-6111, 1990
A:Title: Characterization of the gene for human plasminogen, a key proenzyme in the fibr
A:Reference number: A35229; PMID:90202879; PMID:2318848
A:Accession: A35229
A:Molecule type: DNA
A:Residues: 1-810 <PRT>
A:Cross-references: GB:U05286; GB:M34276; NID:g190064; PIDD:AAA60113.1; PID:g387026
R:Experimental source: leukocyte, lung fibroblast
R:Waldarrett, N.; Bruno, L.; Pontoglio, M.; Candiani, G.; Meroni, G.; Ottolenghi, S.; T
Biochem. Biophys. Res. Commun. 173, 1013-1018, 1990
A:Title: Definition of the transcription initiation site of human plasminogen gene in li
A:Reference number: I52242; PMID:91097553; PMID:2266308
A:Accession: I52242
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-16 <MAL>
A:Cross-references: GB:M62890; NID:g190092; PIDD:AAA36454.1; PID:g553613
R:Forrest, M.; Raden, B.; Israelsson, M.; Larsson, K.; Hedén, L.O.
FEBS Lett. 213, 254-260, 1987
A:Title: Molecular cloning and characterization of a full-length cDNA clone for human pl
A:Reference number: A26646; PMID:87162490; PMID:3030813
A:Accession: A26646
A:Molecule type: mRNA
A:Residues: 1-471; 'D', 473-810 <FOR>
A:Cross-references: GB:X05199; NID:g35530; PIDD:CAA28831.1; PID:g35531
R:Experimental source: liver
R:Malinowski, D.P.; Sadler, J.E.; Davie, E.W.
Biochemistry 23, 4243-4250, 1984
A:Title: Characterization of a complementary deoxyribonucleic acid coding for human and
A:Reference number: I45961; PMID:85023311; PMID:6148961

A:Accession: I62738
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 292-471; 'D', 473-810 <MAL>
A:Cross-references: GB:K02922; NID:g190112; PIDD:AAA60124.1; PID:g387031
A:Accession: I64609
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 367-419 <ML3>
A:Cross-references: GB:K02921; NID:g190110; PIDD:AAA60123.1; PID:g190111
R:Brustholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergler, W.; Manneberg, M
Eur. J. Biochem. 114, 465-470, 1981
A:Title: Comparison of the primary structure of the N-terminal CNBr fragments of human,
A:Reference number: S03335; PMID:8121097; PMID:7238497
A:Accession: S03335
A:Molecule type: protein
A:Residues: 20-71; 'E', 73-76 <BRU>
R:Sortrup-Jensen, L.; Petersen, T.E.; Magnusson, S.
submitted to the Atlas, July 1977
A:Reference number: A00929
A:Accession: A00929
A:Molecule type: protein
A:Residues: 20-71; 'E', 73-85, 87-106, 'D', 108-360, 'E', 362-810 <SOT>
R:Wiman, B.
Eur. J. Biochem. 76, 129-137, 1977
A:Title: Primary structure of the B-chain of human plasmin.
A:Reference number: A04627; PMID:7725245; PMID:142009
A:Accession: A04627
A:Molecule type: protein
A:Residues: 581-810 <WII>
R:Wiman, B.; Wallen, P.
Eur. J. Biochem. 50, 469-494, 1975
A:Title: Structural relationship between "glutamic acid" and "lysine" forms of human pla
A:Reference number: A04625; PMID:75093329; PMID:122932
A:Accession: A04625
A:Molecule type: protein
A:Residues: 20-50; 'Q', 51-71, 'E', 73-85, 87-100 <W12>
R:Wiman, B.; Wallen, P.
Eur. J. Biochem. 58, 539-547, 1975
A:Title: Amino-acid sequence of the cyanogen-bromide fragment from human plasminogen tha
A:Reference number: A04626; PMID:76043692; PMID:128863
A:Accession: A04626
A:Molecule type: Protein
A:Residues: 483-507; 'E', 509-604 <W13>
R:Robbins, K.C.; Bernabe, P.; Arzadon, L.; Summaria, L.
J. Biol. Chem. 248, 1631-1633, 1973
A:Title: The primary structure of human plasminogen. II. The histidine loop of human pl-
A:Reference number: A92125; PMID:73149248; PMID:4694729
A:Contents: annotation; active site
R:Großkopf, W.R.; Summaria, L.; Robbins, K.C.
J. Biol. Chem. 244, 3590-3597, 1969
A:Title: Studies on the active center of human plasmin. Partial amino acid sequence of a
A:Reference number: A92048; PMID:69234739; PMID:4240117
A:Contents: annotation; active site
R:Textler, M.; Vail, Z.; Pecthy, L.
J. Biol. Chem. 257, 7401-7406, 1982
A:Title: Structure of the omega-aminocarboxylic acid-binding sites of human plasminogen.
A:Reference number: A92182; PMID:82213905; PMID:6919539
A:Contents: annotation; omega-aminocarboxylic acid binding sites
R:Vail, Z.; Pecthy, L.
J. Biol. Chem. 259, 13690-13694, 1984
A:Title: The fibrin-binding site of human plasminogen. Arginines 32 and 34 are essential
A:Reference number: A92458; PMID:85054794; PMID:6094526
A:Contents: annotation; fibrin binding site; omega-aminocarboxylic acid binding site
R:Caio, Y.; Uli, R.W.; Davidson, D.; Schaller, J.; Marti, D.; Soehnle, S.; McCance, S.G.;
J. Biol. Chem. 271, 29461-29467, 1996
A:Title: Kringle domains of human angiotensin. Characterization of the anti-proliferativ
A:Reference number: A58811; PMID:97067211; PMID:8910613
A:Contents: annotation
R:Jensen, H.R.; Ugn, F.; Bini, A.; Collier, D.
Biochemistry 37, 4699-4702, 1998
A:Title: Generation of an angiotensin-like fragment from plasminogen by streptolysin-1 (MT

A:Reference number: A58812; MUID:9548733; PMID:9548733
A:Contents: annotation
R:Tulinsky, A.; Mulichak, A.M.
submitted to the Brookhaven Protein Data Bank, July 1991
A:Reference number: A51341; PDB:1PK4
A:Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 376-454
R:Tulinsky, A.; Wu, T.P.
submitted to the Brookhaven Protein Data Bank, July 1991
A:Reference number: A51488; PDB:2PK4
A:Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 375-454
R:Wu, T.P.; Tulinsky, A.
submitted to the Brookhaven Protein Data Bank, August 1993
A:Reference number: A51911; PDB:1PKR
A:Contents: annotation; X-ray crystallography, 2.48 angstroms, residues 102-181
R:Padmanabhan, K.; Tulinsky, A.
submitted to the Brookhaven Protein Data Bank, April 1994
A:Reference number: A52408; PDB:1PMK
A:Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 377-454
R:Tulinsky, A.; Mathews, I.I.
submitted to the Brookhaven Protein Data Bank, December 1995
A:Reference number: A65244; PDB:1CEA
A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
R:Tulinsky, A.; Mathews, I.I.
submitted to the Brookhaven Protein Data Bank, December 1995
A:Reference number: A65245; PDB:1CEB
A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
R:Mulichak, A.M.; Tulinsky, A.; Ravichandran, K.G.
Biochemistry 30, 10576-10588, 1991
A:Title: Crystal and molecular structure of human plasminogen kringle 4 refined at 1.9 Å
A:Reference number: A58819; MUID:92031502; PMID:1657148
A:Contents: annotation
R:Wu, T.P.; Padmanabhan, K.; Tulinsky, A.; Mulichak, A.M.
Biochemistry 30, 10589-10594, 1991
A:Title: The refined structure of the epsilon-aminocaproic acid complex of human plasmin
A:Reference number: A58818; MUID:92031503; PMID:1657149
A:Contents: annotation
R:de Vos, A.M.; Ultsch, M.H.; Kelley, R.F.; Padmanabhan, K.; Tulinsky, A.; Westbrook, M.
Biochemistry 31, 270-279, 1992
A:Title: Crystal structure of the kringle 2 domain of tissue plasminogen activator at 2.4 Å
A:Reference number: A39483; MUID:92118803; PMID:1310033
A:Contents: annotation; X-ray crystallography, 2.4 angstroms
R:Stec, B.; Teeler, M.M.; Whitlow, M.; Yamano, A.
submitted to the Brookhaven Protein Data Bank, June 1995
A:Reference number: A65980; PDB:1KRN
A:Contents: annotation; X-ray crystallography, 1.67 angstroms, residues 376-454
R:Rejzante, M.; Llinas, M.
submitted to the Brookhaven Protein Data Bank, August 1996
A:Reference number: A65803; PDB:1HPJ
A:Contents: annotation; conformation by (1)H-NMR, residues 103-181
R:Rejzante, M.; Llinas, M.
submitted to the Brookhaven Protein Data Bank, August 1996
A:Reference number: A65804; PDB:1HPK
A:Contents: annotation; conformation by (1)H-NMR, residues 103-181
R:Rejzante, M.R.; Llinas, M.
Eur. J. Biochem. 221, 927-937, 1994
A:Title: (1)H-NMR assignments and secondary structure of human plasminogen kringle 1.
A:Reference number: S43645; MUID:94237157; PMID:8181475
A:Contents: annotation; conformation by (1)H-NMR, residues 96-184.
R:Rejzante, M.R.; Llinas, M.
Eur. J. Biochem. 221, 939-949, 1994
A:Title: Solution structure of the epsilon-aminohexanoic acid complex of human plasminogen
A:Reference number: A58817; MUID:94237158; PMID:8181476
A:Contents: annotation; conformation by (1)H-NMR
C:Comment: plasminogen is synthesized by the kidney and is present in plasma and many of
d fibrinogen is converted to plasmin by plasminogen activators (see PIR:UKH0
d PIR:IGHUGB).
C:Comment: Plasmin is inactivated by alpha-2-antiplasmin (see PIR:ITHU02) immediately af
ter-580, resulting in two chains connected by two disulfide bonds. Without the inhibitor,
C:Comment: Microplasmin is formed by autocatalytic cleavage of plasmin under artificial cond
C:Comment: Streptolysin 1 (see PIR:KCHUS1) acts on plasminogen to produce angiotensin. To
ling solid tumors.
C:Genetics:
A:Gene: GDB:PLG

A:Cross-references: GDB:119498; OMIM:173350
A:Map position: 6q26-6q27
A:Introns: 17/1; 62/2; 98/1; 136/2; 183/1; 223/2; 263/1; 317/2; 366/1; 419/2; 480/1; 525
C:Function:
A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a va
n the walls of the graafian follicle; also activates the urokinase-type plasminogen ac
A:Pathway: fibrinolysis
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: angiogenesis inhibitor; blood; duplication; fibrinolysis; glycoprotein; hydr
F:1-96/Domain: plasminogen-related protein predicted <SIG>
F:1-19/Domain: signal sequence #status experimental <PRO>
F:20-810/Product: plasminogen peptide #status experimental <APR>
F:20-96/Domain: activation peptide #status experimental <AST>
F:79-466/Product: angiotensin #status experimental <MAT>
F:97-580,581-810/Product: plasmin #status experimental <CHA>
F:97-580/Domain: plasmin chain A #status experimental <KRI>
F:103-181/Domain: kringle homology <KRI>
F:185-262/Domain: kringle homology <KR2>
F:275-352/Domain: kringle homology <KR3>
F:377-454/Domain: kringle homology <KR4>
F:481-560/Domain: kringle homology <KR5>
F:550-580,581-810/Product: microplasmin #status experimental <MMT>

Query Match 11.0%; Score 142; DB 1; Length 810;
Best Local Similarity 37.2%; Pred. No. 0.00044;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

QY 4 CFWMNGHLYREDQSPAPGRLCLNWLDA-----OSGLASAPVSGAGNHSYCRNPDDEPRG 58
DB 103 CKTGNKGVYGTMSKTKNGITTCQKMSSTSPHRRPSPATIPSGRL-ENNYCRNPDDNDPG 161

QY 59 PWCYSGAGVPEKR--PCEDLRCP 82
DB 162 PWCYTTD-----PEKRYDYCDILBCEB 183

RESULT 7
A40522
plasmin (EC 3.4.21.7) precursor - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 28-Feb-1992 #sequence_revision 17-Apr-1993 #text_change 16-Jul-1999
C:Accession: A40522
R:Kanadas, J.J.; Makker, S.P.
J. Biol. Chem. 266, 10825-10829, 1991
A:Title: Identification of the rat Heymann nephritis autoantigen (GP330) as a receptor
A:Reference number: A40522; MUID:91250378; PMID:1645711
A:Accession: A40522
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-169 <KAN>
A:Cross-references: GB:M62832; NID:9206215; PIDN:AAA41884.1; PID:9554488
A>Note: the authors translated the codon TCG for residue 76 as Ala
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:34-112/Domain: kringle homology <KR6>
F:34-112,55-95,83-107/Disulfide bonds: #status predicted

Query Match 10.7%; Score 138; DB 2; Length 169;
Best Local Similarity 31.8%; Pred. No. 0.00016;
Matches 35; Conservative 13; Mismatches 48; Indels 14; Gaps 5;

QY 4 CFWMNGHLYREDQSPAPGRLCLNWLDA-----OSGLASAPVSGAGNHSYCRNPDDEPRG 58
DB 34 CYQNGKSKYRGTSTWTKGKQSWMTPHSHSKTPANFPDGL-ENNYCRNPDDNDPG 92

QY 59 PWCYSGAGVPEKR--PCEDLRCPETTSQALPAFTTEIQASASGPGADE 106
DB 93 PWCFTTD---PSYRWEXCNLRCSFTGGV--AESALVPOVPSAPGTSE 136

RESULT 8
UKBAY
u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon

F:128-209/Domain: kringle homology <KRG>
F:326-471/Domain: trypsin homology <TRY>
F:42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4
F:272,321,428/Active site: His, Asp, Ser #status predicted

Query Match
Best Local Similarity 10.4%; Score 134; DB 1; Length 477;
Matches 27; Conservative 9; Mismatches 20; Indels 14; Gaps 3;

QY 4 CFWDNGHLYREDQTSAPAGRLCLNMLDAQSGL-----ASAPVSGAGNHSYCRNPD 53
DB 128 CYKDCGVTVRGWSTSESAQCINM---NSNLLTRRTYNGRSDAITLGLGNHNYCRNPD 184
QY 54 EDRPGMPCVY 63
DB 185 NNSK-FMPCVY 193

RESULT 12
J50598
t-plasminogen activator (EC 3.4.21.68) alpha-2 precursor - common vampire bat
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: J50598
R:Kraetzschmar, U.; Haendler, B.; Langer, G.; Boisdol, W.; Bringmann, P.; Alagon, A.; Dot
Gene 105, 223-237, 1991
A:Title: The plasminogen activator family from the salivary gland of the vampire bat Des
A:Reference number: J50597; MUID:92039036; PMID:1937019
A:Accession: J50598
A:Molecule type: mRNA
A:Residues: 1-477 <KRA>
A:Cross-references: GB:M63988; NID:g166074; PID:AAA31593.1; PID:g166075
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: signal sequence #status predicted <SIG>
F:37-477/Product: plasminogen activator alpha-2 #status predicted <PLA>
F:42-79/Domain: fibronectin type I repeat homology <1FA>
F:87-120/Domain: EGF homology <EGF>
F:128-209/Domain: kringle homology <KRG>
F:326-471/Domain: trypsin homology <TRY>
F:42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4
F:185-398/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:225-226/Cleavage site: His-Ser (plasmin) #status predicted
F:272,321,428/Active site: His, Asp, Ser #status predicted

Query Match
Best Local Similarity 10.4%; Score 134; DB 2; Length 477;
Matches 27; Conservative 9; Mismatches 20; Indels 14; Gaps 3;

QY 4 CFWDNGHLYREDQTSAPAGRLCLNMLDAQSGL-----ASAPVSGAGNHSYCRNPD 53
DB 128 CYKDCGVTVRGWSTSESAQCINM---NSNLLTRRTYNGRSDAITLGLGNHNYCRNPD 184
QY 54 EDRPGMPCVY 63
DB 185 NNSK-FMPCVY 193

RESULT 13
J5061
macrophage-stimulating protein 1 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 16-Jun-2000
C:Accession: J5061
R:Ohshiro, K.; Iwama, A.; Matsuo, K.; Ezaki, T.; Sakamoto, O.; Hamaguchi, I.; Takasu, N
Biochem. Biophys. Res. Commun. 227, 273-280, 1996
A:Title: Molecular cloning of rat macrophage-stimulating protein and its involvement in
A:Reference number: J5061; MUID:97011126; PMID:8858136
A:Accession: J5061
A:Molecule type: mRNA
A:Residues: 1-716 <OHS>

A:Cross-references: EMBL:X95096; NID:g1669718; PID:CAA64473.1; PID:g1669719
C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
C:Keywords: duplication; glycoprotein; growth factor; kringle
F:1-31/Domain: signal sequence #status predicted <SIG>
F:32-488,489-716/Product: macrophage-stimulating protein 1 #status predicted <MAT>
F:312-488/Domain: macrophage-stimulating protein 1 alpha chain #status predicted <ACH>
F:110-186/Domain: kringle homology <KR11>
F:191-268/Domain: kringle homology <KR12>
F:292-370/Domain: kringle homology <KR13>
F:379-457/Domain: kringle homology <KR14>
F:488-716/Domain: macrophage-stimulating protein 1 beta chain #status predicted <BCH>
F:489-709/Domain: trypsin homology <TRY>
F:72,305,620/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match
Best Local Similarity 10.4%; Score 133.5; DB 1; Length 716;
Matches 41; Conservative 9; Mismatches 48; Indels 43; Gaps 6;

QY 4 CFWDNGHLYREDQTSAPAGRLCLNMLDAQSGLASAPVSGAGNHSYCRNPDPRGPM 60
DB 110 CIMNGASYSRQTVARTADGLPCQMSRFRPNDRKTYTPKNGL-EEVFCNPDGDRGPM 168
QY 61 CYVS-----GEAGYPER-RPCE--DLRCPEP----- 83
DB 169 CYTTRRSVRFSQSGIKSCREAVCWNGEDYRGSDVDTESGREGQRMWLDGHPHSHPHPE 228
QY 84 --TSQALPAFTTEIQEASDEP 102
DB 229 KFPDKALNDYCRNPDASERP 249

RESULT 14
UKHU
u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human
N:Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasminogen
N:Contains: urokinase-type plasminogen activator chain A; urokinase-type plasminogen act
in form
C:Species: Homo sapiens (man)
C>Date: 17-Dec-1982 #sequence_revision 04-Dec-1986 #text_change 15-Sep-2000
C:Accession: A00931; I52209; JT0102; A37561; I38102; S65783; A37562; A37563; A37564; A35
R:Riccio, A.; Grimaldi, G.; Verde, P.; Sebastio, G.; Boast, S.; Blasf, F.
Nucleic Acids Res. 13, 2759-2771, 1985
A:Title: The human urokinase-Plasminogen activator gene and its promoter.
A:Reference number: A00931; MUID:85215647; PMID:2987867
A:Accession: A00931
A:Molecule type: DNA
A:Residues: 1-431 <RIC>
A:Cross-references: GB:X02419; NID:g37601; PID:CAA26268.1; PID:g1834524
A:Note: the authors translated the codon ATG for residue 214 as Ile
R:Nagamine, Y.; Pearson, D.; Gratian, M.
Biochem. Biophys. Res. Commun. 132, 563-569, 1985
A:Title: Exon-intron boundary sliding in the generation of two mRNAs coding for porcine
A:Reference number: I52209; MUID:86050639; PMID:393505
A:Accession: I52209
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 145-161 <NAG1>
A:Cross-references: GB:X03027; NID:g340174; PID:AAA61257.1; PID:g340175
R:Nagai, M.; Hiramatsu, R.; Kaneda, T.; Hayasuke, N.; Arimura, H.; Nishida, M.; Suyama,
Gene 36, 183-188, 1985
A:Title: Molecular cloning of cDNA coding for human preprourokinase.
A:Reference number: JT0102; MUID:86056954; PMID:2415429
A:Accession: JT0102
A:Molecule type: mRNA
A:Residues: 1-213,1,215-431 <NAG2>
A:Cross-references: GB:X03226; NID:g340155; PID:AA97138.1; PID:g340158; GB:D00244; NID
R:Verde, P.; Stoppelli, M.P.; Galeffi, P.; Di Nocera, P.; Blasi, F.
Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984
A:Title: Identification and primary structure of an unspliced human urokinase poly(A) + RN
A:Reference number: A37561; MUID:84727206; PMID:6589620
A:Accession: A37561
A:Molecule type: mRNA

Mon Apr 7 10:22:03 2003

us-10-057-951-2_copy_22_263.rpr

Page 9

[illegible]

Search completed: April 7, 2003, 09:19:19
Job time : 21.6216 secs

GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:24 ; Search time 11.0686 Seconds

(without alignments)
906.823 Million cell updates/sec

Title: US-10-057-951-2_COPY_22_263

Sequence: 1 SGGCFWMDNGLHYREDQTPA.....PVDPOEGSTPLMGQAQTRGA 242

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	156.5	12.1	566	1	TPA BOVIN
2	154.5	12.0	562	1	TPA HUMAN
3	154	11.9	559	1	TPA RAT
4	150	11.6	653	1	HGF α MOUSE
5	147.5	11.4	559	1	TPA MOUSE
6	146.5	11.4	655	1	HGF α HUMAN
7	142	11.0	810	1	PLMN HUMAN
8	138	10.7	169	1	PLMN RAT
9	138	10.7	433	1	UROK PAPCY
10	135.5	10.5	442	1	UROK PIG
11	134	10.4	431	1	URTB DESRO
12	134	10.4	477	1	URTB DESRO
13	132	10.2	431	1	URTB DESRO
14	130.5	10.1	716	1	HGF α MOUSE
15	129	10.0	810	1	PLMN ERIBU
16	127	9.9	394	1	URTB DESRO
17	127	9.9	433	1	UROK BOVIN
18	126.5	9.8	728	1	HGF MOUSE
19	124.5	9.7	810	1	PLMN MACMU
20	124	9.6	622	1	THRB HUMAN
21	124	9.6	728	1	HGF HUMAN
22	124	9.6	812	1	PLMN BOVIN
23	123.5	9.5	434	1	UROK CHICK
24	123	9.5	593	1	FA12 BOVIN
25	122.5	9.5	4548	1	APOA HUMAN
26	122	9.5	603	1	FA12 CAURO
27	121.5	9.4	728	1	HGF RAT
28	120.5	9.3	333	1	PLMN CANRA
29	120	9.3	790	1	PLMN PIG
30	119.5	9.3	477	1	URTB DESRO
31	119	9.2	615	1	UROK MOUSE
32	118.5	9.2	433	1	FA12 HUMAN
33	118	9.2	432	1	UROK RAT

34	117	9.1	625	1	THRB BOVIN	P00735 bos taurus
35	116	9.0	1420	1	APOA MACMU	P14417 macaca mula
36	115.5	9.0	711	1	HGF α HUMAN	P26927 homo sapien
37	114.5	8.9	812	1	PLMN MOUSE	P20918 mus musculu
38	114	8.8	473	1	KREM MOUSE	Q99n43 mus musculu
39	114	8.8	473	1	KREM RAT	Q92464 ratus norv
40	114	8.8	475	1	KREM HUMAN	Q96mu8 homo sapien
41	114	8.8	618	1	THRB MOUSE	P19221 mus musculu
42	109.5	8.5	1709	1	SN HUMAN	Q9b222 homo sapien
43	107	8.3	617	1	THRB RAT	P18292 ratus norv
44	103.5	8.0	343	1	PLMN SHEEP	P81286 ovis aries
45	101.5	7.9	325	1	PLMN PETMA	P33574 petromyzon

ALIGNMENTS

```

RESULT 1
ID   TPA_BOVIN
AC   Q26198;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE   (t-PA) (t-plasminogen activator).
GN   tPA.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OX   Bovidae; Bovinae; Bos.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Ravn P., Berglund L., Petersen T.E.;
RT   "Cloning and characterization of the bovine plasminogen activators uPA
RT   and tPA."
RL   Int. Dairy J. 5:605-617(1995).
CC   -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN
CC   TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC   CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC   ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC   MANY OTHER PHYSIOPATHOLOGICAL EVENTS.
CC   -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC   plasminogen to form plasmin.
CC   -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC   BOND.
CC   -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC   -1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC   PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC   ARG-314 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC   -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC   CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC   -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC   -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC   -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC   -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC   THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; X85800; CAA59795.1; -.
DR   HSSP; P00750; IRTF.
DR   MEROPS; S01.232; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR000083; Fibrinctn1.

```

```
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00039; fn1_1.
DR Pfam: PF00051; kringle_2.
DR Pfam: PF00089; trypsin_1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR0018; KRINGLE.
DR ProDom: PD000395; Kringle_2.
DR SMART: SM00181; EGF_1.
DR SMART: SM00058; FN1_1.
DR SMART: SM00130; KR_2.
DR SMART: SM00020; TRYD_SPC_1.
DR PROSITE: PS00022; EGF_1.
DR PROSITE: PS0186; EGF_2_1.
DR PROSITE: PS01253; FIBRONECTIN_1_1.
DR PROSITE: PS00021; KRINGLE_1_1.
DR PROSITE: PS50070; TRYPsin_DOM_1.
DR PROSITE: PS50240; TRYPsin_DOM_1.
DR PROSITE: PS00134; TRYPsin_HIS_1.
DR PROSITE: PS00135; TRYPsin_SER_1.
KM Plasmogen activation; Hydrolase; Serine protease; Glycoprotein;
FT SIGNAL 1 21
FT PROPEP 22 33
FT CHAIN 34 566
FT CHAIN 34 314
FT CHAIN 315 566
FT DOMAIN 40 82
FT DOMAIN 83 121
FT DOMAIN 128 209
FT DOMAIN 219 300
FT DOMAIN 315 566
FT ACT_SITE 361 361
FT ACT_SITE 410 410
FT ACT_SITE 517 517
FT ACT_SITE 42 72
FT DISULFID 70 79
FT DISULFID 87 98
FT DISULFID 92 109
FT DISULFID 111 120
FT DISULFID 128 209
FT DISULFID 149 191
FT DISULFID 180 204
FT DISULFID 219 300
FT DISULFID 240 282
FT DISULFID 271 295
FT DISULFID 303 434
FT DISULFID 346 362
FT DISULFID 354 423
FT DISULFID 448 523
FT DISULFID 480 496
FT DISULFID 513 541
FT CARBOHYD 153 153
FT CARBOHYD 487 487
SQ SEQUENCE 566 AA; 63701 MW; 2EB6BB4B32276C3 CRC64;

Query Match 12.1%; Score 156.5; DB 1; Length 566;
Best Local Similarity 36.7%; Pred. No. 6.4e-06;
Matches 36; Conservative 10; Mismatches 39; Indels 13; Gaps 4;
```

```
TPA_HUMAN
ID TPA_HUMAN STANDARD; PRT; 562 AA.
AC P00750; Q15103;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator) (Alteplase) (Retepase).
GN PLAT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Melanoma;
RX MEDLINE=83115262; PubMed=6337343;
RA Pennica D., Holmes W.E., Kohr W.J., Harkins R.N., Vohar G.A., H.L.,
RA Ward C.A., Bennett W.F., Yelverton E., Seeburg P.H., Heyneker H.L.,
RA Goeddel D.V., Collen D.;
RT "Cloning and expression of human tissue-type plasminogen activator
RT cDNA in E. coli."
RL Nature 301:214-221 (1983).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Fetal lung;
RX MEDLINE=88262579; PubMed=3133640;
RA Sasaki H., Saito Y., Hayaishi M., Otsuka K., Niwa M.;
RT "Nucleotide sequence of the tissue-type plasminogen activator cDNA
RT from human fetal lung cells."
RL Nucleic Acids Res. 16:5695-5695 (1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88054470; PubMed=2824147;
RA Reddy V.B., Garramone A.J., Sasek H., Wei C.-M., Watkins P., Galli J.,
RA Hsling N.;
RT "Expression of human uterine tissue-type plasminogen activator in
RT mouse cells using BPV vectors."
RL DNA 6:461-472 (1987).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=86196143; PubMed=3009482;
RA Flierzer Degen S.O., Rajput B., Reich E.;
RT "The human tissue plasminogen activator gene."
RL J. Biol. Chem. 261:6972-6985 (1986).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=84298137; PubMed=6089198;
RA Ny T., Elgh F., Lund B.;
RT "The structure of the human tissue-type plasminogen activator gene:
RT correlation of intron and exon structures to functional and
RT structural domains."
RL Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359 (1984).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=86284200; PubMed=3090401;
RA Harris T.J., Patel T., Marston F.A., Little S., Emtage J.S.,
RA Odenakker G., Volckert G., Rombaux W., Balliau A., Somer P.;
RT "Cloning of cDNA coding for human tissue-type plasminogen activator
RT and its expression in Escherichia coli."
RL Mol. Biol. Med. 3:279-292 (1986).
RN [7]
RP SEQUENCE OF 212-361 FROM N.A.
RX MEDLINE=83169656; PubMed=6572897;
RA Edlund T., Ny T., Raahy M., Heden L.-O., Palm G., Holmgren E.,
RA Josephson S.;
RT "Isolation of cDNA sequences coding for a part of human tissue
RT plasminogen activator."
RL Proc. Natl. Acad. Sci. U.S.A. 80:349-352 (1983).
RN [8]
RP SEQUENCE OF 1-36 FROM N.A.
RX MEDLINE=85289338; PubMed=3161893;
RA Fisher R., Waller E.K., Grossi G., Thompson D., Tizard R.,
```


RA Schlemming W.-D.;
RT "Isolation and characterization of the human tissue-type plasminogen
RT activator structural gene including its 5' flanking region.";
RL J. Biol. Chem. 260:11223-11230(1985).
RN [9]
RP SEQUENCE OF 31-562 FROM N.A.
RX MEDLINE=91291340; PubMed=636681;
RA Itagaki Y., Yasuda H., Motiura T., Mitsuda S., Higashio K.;
RT "Purification and characterization of tissue plasminogen activator
RT secreted by human embryonic lung diploid fibroblasts, IMR-90 cells";
RL Agric. Biol. Chem. 55:1225-1232(1991).
RN [10]
RP SEQUENCE OF 36-562.
RC TISSUE=Melanoma;
RX MEDLINE=85000468; PubMed=6433976;
RA Pohl G., Kaelin-Stroem M., Bergsdorf N., Wallen P., Joernvall H.;
RT "Tissue plasminogen activator: peptide analyses confirm an indirectly
RT derived amino acid sequence, identify the active site serine residue,
RT establish glycosylation sites, and localize variant differences.";
RL Biochemistry 23:3701-3707(1984).
RN [11]
RP SEQUENCE OF 33-52 AND 311-330.
RC TISSUE=Melanoma;
RX MEDLINE=83209620; PubMed=6682760;
RA Wallen P., Pohl G., Bergsdorf N., Raanby M., Ny T., Joernvall H.;
RT "Purification and characterization of a melanoma cell plasminogen
RT activator";
RL Eur. J. Biochem. 132:681-686(1983).
RN [12]
RP SEQUENCE FROM N.A. (SHORT ISOFORM).
RC TISSUE=Umbilical vein;
RX MEDLINE=90192129; PubMed=2107528;
RA Siebert P.D., Fong K.;
RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from
RT human endothelial cells";
RL Nucleic Acids Res. 18:1086-1086(1990).
RN [13]
RP STRUCTURE OF CARBOHYDRATES.
RX MEDLINE=90092112; PubMed=2513186;
RA Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.;
RT "Carbohydrate structure of recombinant human uterine tissue
RT plasminogen activator expressed in mouse epithelial cells";
RL Eur. J. Biochem. 186:273-286(1989).
RN [14]
RP CARBOHYDRATE-LINKAGE SITE THR-96.
RX MEDLINE=91159408; PubMed=1900431;
RA Harris R.U., Leonard C.K., Guzzetta A.W., Spellman M.W.;
RT "Tissue plasminogen activator has an O-linked fucose attached to
RT chondrocyte-61 in the epidermal growth factor domain";
RL Biochemistry 30:2311-2314(1991).
RN [15]
RP DISULFIDE BONDS IN KRINGLE 2.
RX MEDLINE=91244765; PubMed=1645336;
RA Vlahos C.U., Wilhelm O.G., Hassell T., Jaskunas S.R., Bang N.U.;
RT "Disulfide pairing of the recombinant kringle-2 domain of tissue
RT plasminogen activator produced in Escherichia coli.";
RL J. Biol. Chem. 266:10070-10072(1991).
RN [16]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF CATALYTIC DOMAIN.
RX MEDLINE=96200985; PubMed=8613982;
RA Lamba D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,
RA Bode W.;
RT "The 2.3 A crystal structure of the catalytic domain of recombinant
RT two-chain human tissue-type plasminogen activator";
RL J. Mol. Biol. 258:117-135(1996).
RN [17]
RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.
RX MEDLINE=97449126; PubMed=9305622;
RA Renatus M., Engh R.A., Stubbs M.T., Huber R., Fischer S., Kohnert U.,
RA Bode W.;
RT "Lysine 156 promotes the anomalous proenzyme activity of tPA: X-ray
RT crystal structure of single-chain human tPA.";
RL EMBO J. 16:4797-4805(1997).

```

RN [18] X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF KRINGLE 2.
RX MEDLINE=92118803; Pubmed=1310033;
RA de Vos A., Ulsch M.H., Kelley R.F., Padmanabhan K., Tullinsky A.,
RA Westbrock M.L., Kosciakof A.A.;
RT "Cyteal structure of the kringle 2 domain of tissue plasminogen
RT activator at 2.4-A resolution."
RL Biochemistry 31:270-279(1992).
RN [19]
RP STRUCTURE BY NMR OF KRINGLE 2.
RX MEDLINE=90122799; Pubmed=2558718;
RA Byeon I.-J.L., Kelley R.F., Llinas M.;
RT "H NMR structural characterization of a recombinant kringle 2 domain
RT from human tissue-type plasminogen activator."
RL Biochemistry 28:9350-9360(1989).
RN [20]
RP STRUCTURE BY NMR OF KRINGLE 2.
RX MEDLINE=91200042; Pubmed=1901789;
RA Byeon I.-J.L., Kelley R.F., Llinas M.;
RT "Kringe-2 domain of the tissue-type plasminogen activator. 1H-NMR
RT assignments and secondary structure."
RL Eur. J. Biochem. 197:155-165(1991).
RN [21]
RP STRUCTURE BY NMR OF KRINGLE 2.
RX MEDLINE=92106329; Pubmed=1762144;
RA Byeon I.-J.L., Llinas M.;
RT "Solution structure of the tissue-type plasminogen activator kringle
RT 2 domain complexed to 6-aminohexanoic acid an antifibrinolytic
RT drug."
RL J. Mol. Biol. 222:1035-1051(1991).
RN [22]
RP STRUCTURE BY NMR OF 38-85.
RX MEDLINE=92292163; Pubmed=1602484;
RA Downing A.K., Driscoll P.C., Harvey T.S., Dudgeon T.J., Smith B.O.,
RA Baron M., Campbell I.D.;
RT "Solution structure of the fibrin binding finger domain of
RT tissue-type plasminogen activator determined by 1H nuclear magnetic
RT resonance."
RL J. Mol. Biol. 225:821-833(1992).
RN [23]
RP STRUCTURE BY NMR OF 36-126.
RX MEDLINE=96027104; Pubmed=7582893;
RA Smith B.O., Downing A.K., Driscoll P.C., Dudgeon T.J., Campbell I.D.;
RT "The solution structure and backbone dynamics of the fibronectin type
RT I and epidermal growth factor-like pair of modules of tissue-type
RT plasminogen activator."
RL Structure 3:823-833(1995)
CC -I- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN
CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC MANY OTHER PHYSIOPATHOLOGICAL EVENTS.
CC -I- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -I- SUBUNIT: BINDS TO FIBRIN WITH HIGH AFFINITY. THIS INTERACTION
DB Query Match 12.0%; Score 154.5; DB 1; Length 562;
Beat Local Similarity 39.6%; Pred. No. 9.4e-06;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
QY 4 CFWDNGHLVREDQTSPPAGLRCLNWLDAOSGLASPVS-----GAGNSHYCRNPDE 54
DB 127 CYEDQGISYRWGTVAESAEGECTNW--NSSALAOKPYSGRPRDAIRLIGLNHNYYCRNPDR 184
QY 55 DRGPWCYSGRAGVPEKRPCEDLACPERTS 85
DB 185 DSK-PWCVTF-KAGKYSEFCSTPACSEGNs 213

```

DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (BC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator).
GN PLAT.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89170114; PubMed=3148445;
RA NY T., Leonardsson G., Hsueh A.J.W.;
RT "Cloning and characterization of a cDNA for rat tissue-type
RT plasminogen activator.";
RL DNA 7:671-677(1988).
RL [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90130448; PubMed=2105315;
RA Feng P., Ohlsson M., Ny T.;
RT "The structure of the TATA-box rat tissue-type plasminogen activator
RT gene. Species-specific sequence divergences in the promoter predict
RT differences in regulation of gene expression.";
RL J. Biol. Chem. 265:2022-2027(1990).
CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN
CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC MANY OTHER PHYSIOLOGICAL EVENTS.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC BOND.
CC -1- SUBCELLULAR LOCATION: SECRETED, EXTRACELLULAR.
CC -1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALIKREIN OR FACTOR XA.
CC -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, M23697, AAA41812.1, -;
DR EMBL, M31197, AAA442261.1, -;
DR EMBL, M31185, AAA442261.1, JOINED.
DR EMBL, M31186, AAA442261.1, JOINED.
DR EMBL, M31187, AAA442261.1, JOINED.
DR EMBL, M31188, AAA442261.1, JOINED.
DR EMBL, M31189, AAA442261.1, JOINED.
DR EMBL, M31190, AAA442261.1, JOINED.
DR EMBL, M31191, AAA442261.1, JOINED.
DR EMBL, M31192, AAA442261.1, JOINED.
DR EMBL, M31193, AAA442261.1, JOINED.
DR EMBL, M31194, AAA442261.1, JOINED.
DR EMBL, M31195, AAA442261.1, JOINED.
DR EMBL, M31196, AAA442261.1, JOINED.
DR EMBL, A19618, CAA01482.1, -;
DR PIR, A33597, A33597.
DR PIR, A35029, A35029.
DR HSSP, P00750, 1RTF.
DR MEROPS, S01.232, -;
DR InterPro, IPR001314, Chymotrypsin.

DR InterPro, IPR000561, EGF-like.
DR InterPro, IPR000083, Fibrinctn1.
DR InterPro, IPR000001, Kringle.
DR InterPro, IPR001254, Ser_protease_Try.
DR Pfam, PF00008, EGF, 1.
DR Pfam, PF00039, fn1, 1.
DR Pfam, PF00051, kringle, 2.
DR Pfam, PF00089, trypsin, 1.
DR PRINTS, PR00018, KRINGLE.
DR PRODOM, PD000395, Kringle, 2.
DR SMART, SM00181, EGF, 1.
DR SMART, SM00058, EN1, 1.
DR SMART, SM00130, KR, 2.
DR SMART, SM00020, TYP_SPC, 1.
DR PROSITE, PS00022, EGF_2, 1.
DR PROSITE, PS01186, EGF_2, 1.
DR PROSITE, PS01253, FIBRONECTIN_1, 1.
DR PROSITE, PS00021, KRINGLE_1, 2.
DR PROSITE, PS00070, TRYPSIN_DOM, 2.
DR PROSITE, PS00240, TRYPSIN_DOM, 1.
DR PROSITE, PS00134, TRYPSIN_HIS, 1.
DR PROSITE, PS00135, TRYPSIN_SER, 1.
KW plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 17 PROBABLE.
FT PROPE 18 29
FT CHAIN 30 559
FT CHAIN 30 308
FT CHAIN 309 559
FT DOMAIN 36 78
FT DOMAIN 79 117
FT DOMAIN 124 205
FT DOMAIN 213 294
FT DOMAIN 309 559
FT ACT_SITE 355 355
FT ACT_SITE 404 404
FT ACT_SITE 510 510
FT ACT_SITE 510 510
FT DISULFID 38 68
FT DISULFID 66 75
FT DISULFID 83 94
FT DISULFID 88 105
FT DISULFID 107 116
FT DISULFID 124 205
FT DISULFID 145 187
FT DISULFID 176 200
FT DISULFID 213 294
FT DISULFID 234 276
FT DISULFID 265 289
FT DISULFID 297 428
FT DISULFID 340 356
FT DISULFID 348 417
FT DISULFID 442 516
FT DISULFID 474 490
FT DISULFID 506 534
FT CARBOHYD 149 149
FT CARBOHYD 481 481
FT CONFLICT 380 380
SQ SEQUENCE 559 AA; 62903 MM; 7DBD3809C1D1C921 CRC64;
Query Match 11.9%; Score 154; DB 1; Length 559;
Best local Similarity 32.4%; Pred. No. 1e-05;
Matches 45; Conservative 14; Mismatches 54; Indels 26; Gaps 6;
QY 4 CFWDNGHLVREDOTSPAPGRCLNMLDQSGLAGAPVS-----GAGNHSYCRUPDE 54
DB 124 CFEQGITVYRTGWSTAENGACIMW--NSSALSOKPYSARRPNAIKIGLGNHNYCRPPDR 181
QY 55 DPRGWCYVSGEAGVPEKRPCELDRCPE-----TTSQALPAPFTTEIOEASGPG 103
DB 182 DVK-FWCYVF-KAGKYTFEFCTPACPKGPTDCYGVKGYVTRGTHSFTT--SKASCLPW 237

Qy 104 ADEVQVAPANALPARSEA 122
Db 238 NSMILIGKTYTAMRANSDA 256

RESULT 4
HGFA_MOUSE STANDARD; PRT; 653 AA.

AC Q9R098; Q9R098; 40, Created
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF activator) (HGFA).
GN HGFA_MOUSE (Mouse).
OC Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C.
RA Itoh H., Kataoka H., Kono H.;
RT "Mouse hepatocyte growth factor activator."
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RA Yang J., Huan Y.S., Sehgal S., Kukes A., Brady C., Barasch J.,
RT "Activation of HGF by endogenous HGF activator is required for
RT metanephric kidney morphogenesis in vitro."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC - FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY CONVERTING
CC IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM (BY SIMILARITY).
CC - SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A
CC DISULFIDE BOND (BY SIMILARITY).
CC - SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN
CC PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM (BY
CC SIMILARITY).
CC - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC - SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC - SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE II DOMAIN.
CC - SIMILARITY: CONTAINS 1 KIRINGE DOMAIN.
CC - SIMILARITY: CONTAINS 1 KIRINGE DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR EMBL; AF039017; AAF02489.1; -
DR EMBL; AF224724; AAF34712.1; -
DR HSRP; P00763; JDEO.
DR MEROPS; S01.228; -
DR MGD; MGI:1859281; Hgfac.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR000562; FN Type II.
DR InterPro; IPR000083; Fibrinctn.
DR InterPro; IPR000001; Kringie.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF000039; EGF; 2.
DR Pfam; PF000039; fn1; 1.
DR Pfam; PF000040; fn2; 1.
DR Pfam; PF000051; kringie; 1.
DR Pfam; PF000089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.

DR PRINTS; PR00018; KIRINGE.
DR ProDom; PD000395; Kringie; 1.
DR ProDom; PD000995; FN_Type_II; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KIRINGE_1; 1.
DR PROSITE; PS00070; KIRINGE_2; 1.
DR PROSITE; PS00240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolyase; Glycoprotein; Plasma; Serine protease; Kringie; Signal;
KW EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 29
FT PROPEP 30 369
FT CHAIN 370 405
FT CHAIN 406 653
FT DOMAIN 105 145
FT DOMAIN 157 195
FT DOMAIN 197 237
FT DOMAIN 238 276
FT DOMAIN 283 364
FT DOMAIN 406 653
FT ACT_SITE 445 445
FT ACT_SITE 495 495
FT ACT_SITE 596 596
FT ACT_SITE 105 130
FT DISULFID 119 145
FT DISULFID 161 172
FT DISULFID 166 183
FT DISULFID 185 194
FT DISULFID 199 227
FT DISULFID 225 234
FT DISULFID 242 253
FT DISULFID 247 264
FT DISULFID 266 275
FT DISULFID 283 364
FT DISULFID 304 346
FT DISULFID 335 359
FT DISULFID 392 519
FT DISULFID 430 446
FT DISULFID 438 508
FT DISULFID 533 602
FT DISULFID 565 581
FT DISULFID 592 620
FT CARBOHYD 39 39
FT CARBOHYD 47 47
FT CARBOHYD 63 63
FT CARBOHYD 287 287
FT CARBOHYD 466 466
FT CARBOHYD 544 544
FT CONFLICT 164 164
SQ SEQUENCE 653 AA; 70567 MW; 88B4B2025DF7DC CRC64;

Query Match 11.6%; Score 150; DB 1; Length 653;
Best Local Similarity 32.6%; Pred. No. 2.7e-05;
Matches 46; Conservative 17; Mismatches 46; Indels 32; Gaps 7;

Qy 4 CFWDNGHLYREDQTSAPAGLCLNW-----LDAQSGLASAPVSGAGNHSYCRVPD 55
Db 283 CFLNGTEYRGVASTAAGLSCLAMNSDLYQELHVDV-VAAAVLLGLGPHAVCRPDD 341

Qy 56 PRGPNCTYSGENGAVPEKPCEDLACPETTSQALPAFTTEIQEASBEGADDEVQVAPANA 115
Db 342 ER-PWCYVVKDNLASWE-----YCRLLACESLARVHSQTP-----TLA--A 380

QY 116 LPARSEAAVPGISQVR 136
DB 381 LP--ESAPAVRPTCGKHKR 399

RESULT 5
TPA_MOUSE STANDARD; PRT; 559 AA.

AC P1214;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator).
GN PLAT.

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88087303; PubMed=2826484;
RA Rickles R.J., Darrow A.L., Strickland S.;
RT "Molecular cloning of complementary DNA to mouse tissue plasminogen
RT activator mRNA and its expression during F9 teratocarcinoma cell
RT differentiation."
RL J. Biol. Chem. 263:1563-1569(1988).

CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN
CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC MANY OTHER PHYSIOLOGICAL EVENTS.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC BOND.
CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-308 CATALYZED BY PLASMIN. TISSUE KALLIKREIN OR FIBRIN XA.
CC -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J03520; AAA40470.1; -.
DR PIR; A29941; A29941.
DR HSSP; P00750; IASH.
DR MEROPS; S01.232; -.
DR MGD; MGI:97610; Plat.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000083; Fibrinctn.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.

DR SMART; SM00181; EGF_1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 3.
DR PROSITE; PS0070; KRINGLE_2; 2.
DR PROSITE; PS0240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_His; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Plasma; Kringle; EGF-like domain; Hydrolyase; Serine protease; Glycoprotein;
KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 17
FT PROPEP 18 29
FT CHAIN 30 559
FT CHAIN 30 508

FT CHAIN 309 559
FT DOMAIN 36 78
FT DOMAIN 79 117
FT DOMAIN 124 205
FT DOMAIN 213 294
FT DOMAIN 309 559
FT ACT SITE 355 355
FT ACT SITE 404 404
FT ACT SITE 404 404
FT ACT SITE 510 510
FT DISULFID 38 68
FT DISULFID 66 75
FT DISULFID 83 94
FT DISULFID 88 105
FT DISULFID 107 116
FT DISULFID 124 205
FT DISULFID 145 187
FT DISULFID 176 200
FT DISULFID 213 294
FT DISULFID 234 276
FT DISULFID 265 289
FT DISULFID 297 428
FT DISULFID 340 356
FT DISULFID 348 417
FT DISULFID 442 516
FT DISULFID 474 490
FT DISULFID 506 534
FT CARBOHYD 149 149
FT CARBOHYD 481 481

SQ SEQUENCE 559 AA; 63110 MW; 4ACEB57DC6A282A5 CRC64;
Query Match 11.4%; Score 147.5; DB 1; Length 559;
Best Local Similarity 37.0%; Pred. No. 3.7e-05;
Matches 34; Conservative 11; Mismatches 34; Indels 13; Gaps 4;

QY 4 CFMDNGHLYREQGISPARGLCLNWLDAQSLASAPV-----GAGNHSYCRNPDE 54
DB 124 CFEEQGITRGWSTASGAECLNW--NSSVLSLKPYNRRPAIKLGLGNHYCRNPDR 181

QY 55 DRGPWCYVSGEAVPEKRPCEIDRCPEETSQ 86
DB 182 DLK-PWCYVF-KAGKYTFEFCSTPACPKGSE 211

RESULT 6
HGFA_HUMAN STANDARD; PRT; 655 AA.
ID HGFA_HUMAN
AC Q04756; Q14726;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF
DE activator) (HGFA).
GN HGFA.

OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 ON NCI_TaxId=9606;
 [1]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RC TISSUE=Liver, and Serum;
 KK MEDLINE=932878; Pubmed=7683665;
 RA Miyazawa K., Shimomura T., Kitamura A., Kondo J., Morimoto Y.,
 RA Kitamura N.;
 RT "Molecular cloning and sequence analysis of the cDNA for a human
 RT serine protease responsible for activation of hepatocyte growth
 RT factor. Structural similarity of the protease precursor to blood
 RT coagulation factor XII.";
 RL J. Biol. Chem. 268:10024-10028(1993).
 RN [2]
 RP SEQUENCE OF 40-655 FROM N.A.
 RA Zhao S., Odell C.;
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY
 CC CONVERTING IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM.
 CC -1- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A
 CC DISULFIDE BOND.
 CC -1- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN
 CC PROCURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM.
 CC -1- TISSUE SPECIFICITY: LIVER.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE II DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 IS THE INITIATOR.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).

 CC
 DR EMBL: D14012; BAA03113.1; -
 DR EMBL: Z69923; CAA93803.1; -
 DR PIR: A46688; A46688.
 DR HSSP: P00763; IDPO.
 DR MEROPS: S01.228; -
 DR HSP: HSCNC:4894; HGFAC.
 DR MIM: 604552; -
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000742; EGF 2.
 DR InterPro: IPR000562; FN Type II.
 DR InterPro: IPR000083; p1fnctnl.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam: PF00008; EGF 2.
 DR Pfam: PF00039; fn1; 1.
 DR Pfam: PF00040; fn2; 1.
 DR Pfam: PF00051; kringle; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00013; ENTYREIT.
 DR PRINTS: PR00019; KRINGLE.
 DR PRODOM: PD000395; Kringle; 1.
 DR PRODOM: PD000995; FN_Type_II; 1.
 DR SMART: SM00181; EGF; 2.
 DR SMART: SM00058; FN1; 1.
 DR SMART: SM00059; FN2; 1.
 DR SMART: SM00130; KR; 1.
 DR SMART: SM00020; TRYPSIN; 1.
 DR PROSITE: PS00023; EGF 1; 2.
 DR PROSITE: PS01166; EGF 2; 1.
 DR PROSITE: PS01253; FIBRONECTIN_1; 1.

DR PROSITE: PS00023; FIBRONECTIN_2; 1.
 DR PROSITE: PS00021; KRINGLE 1; 1.
 DR PROSITE: PS00070; KRINGLE 2; 1.
 DR PROSITE: PS00240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KM Hydrolyase, Glycoprotein; Plasma, Serine protease; Kringle; Signal;
 KM EGF-like domain; Repeat; Zymogen.
 FT SIGNAL 1 30
 FT PROPEP 31 372
 FT CHAIN 373 407
 FT CHAIN 408 655
 FT DOMAIN 168 198
 FT DOMAIN 160 198
 FT DOMAIN 200 240
 FT DOMAIN 241 279
 FT DOMAIN 286 367
 FT DOMAIN 408 655
 FT ACT SITE 447 447
 FT ACT SITE 497 497
 FT ACT SITE 598 598
 FT ACT SITE 108 133
 FT DISULFID 122 148
 FT DISULFID 164 175
 FT DISULFID 169 186
 FT DISULFID 188 197
 FT DISULFID 202 230
 FT DISULFID 228 237
 FT DISULFID 245 256
 FT DISULFID 250 267
 FT DISULFID 269 278
 FT DISULFID 266 367
 FT DISULFID 307 349
 FT DISULFID 338 362
 FT DISULFID 394 521
 FT DISULFID 432 448
 FT DISULFID 440 510
 FT DISULFID 535 604
 FT DISULFID 567 583
 FT DISULFID 594 622
 FT CARBOHYD 48 48
 FT CARBOHYD 290 290
 FT CARBOHYD 468 468
 FT CARBOHYD 492 492
 FT CARBOHYD 546 546
 FT CONFLICT 644 644
 SQ SEQUENCE 655 AA; 70681 MW; 2CF72F1B1862BD7 CRC64;
 Query Match 11.4%; Score 146.5; DB 1; Length 655;
 Best Local Similarity 36.9%; Pred. No. 5,3e-05;
 Matches 41; Conservative 10; Mismatches 41; Indels 19; Gaps 6;
 Oy 4 CFWDNGHLTYREDOTSPAPGLRLNW-----IDAOSGLASAPVSGAGNHSYCRPPDD 55
 Db 286 CFTGNGYGRVAVSTASGLSCLANNSDLVQELHVDV-VGAALALGLGHAYCRPPDD 344
 Oy 56 PRGFWCVSGEAGVP---EKRPCEDLRCPEPTTSQALPAFTTEIOE-ASRG 101
 Db 345 ER-FWCVVVDLSALSWCYRLACESTL---TRVQLSPDLATLPBPASPG 390
 RESULT 7
 ID PLIN_HUMAN STANDARD; PRT; 810 AA.
 AC P00747;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
 GN PLG
 OS Homo sapiens (human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OK NCBI_TaxId=9606;
RN [11]
RP SEQUENCE FROM N.A.
RX MEDLINE=90202879; PubMed=2318848;
RA Petersen T.E., Martzen M.R., Ichinose A., Davie E.W.;
RT "Characterization of the gene for human plasminogen, a key proenzyme
RT in the fibrinolytic system.";
RL J. Biol. Chem. 265:6104-6111(1990).
RN [12]
RP SEQUENCE FROM N.A.
RX MEDLINE=87162490; PubMed=1030813;
RA Forsgren M., Raden B., Ivarsson M., Larsson K., Heden L.-O.;
RT "Molecular cloning and characterization of a full-length cDNA clone
RT for human plasminogen.";
RL FEBS Lett. 213:254-260(1987).
RN [13]
RP SEQUENCE OF 20-810.
RA Sottrup-Jensen L., Petersen T.E., Magnusson S.;
RL Submitted (JUL-1977) to the EMBL data bank.
RN [14]
RP SEQUENCE OF 292-810 FROM N.A.
RX MEDLINE=85023311; PubMed=6148961;
RA Malinowski D.P., Sadler J.E., Davie E.W.;
RT "Characterization of a complementary deoxyribonucleic acid coding for
RT human and bovine plasminogen.";
RL Biochemistry 23:4243-4250(1984).
RN [15]
RP SEQUENCE OF 20-100.
RX MEDLINE=75093329; PubMed=122932;
RA Wiman B., Wallen P.;
RT "Structural relationship between 'glutamic acid' and 'lysine' forms
RT of human plasminogen and their interaction with the NH2-terminal
RT activation peptide as studied by affinity chromatography.";
RL Eur. J. Biochem. 50:489-494(1975).
RN [16]
RP SEQUENCE OF 95-580; 581-626; 657-700 AND 732-810.
RA Sottrup-Jensen L., Claessens H., Zafdel M., Petersen T.E., Magnusson S.;
RL (in) Davidson J.F., Roman R.M., Samama M.M., Desnoyers P.C. (eds.);
RT Progress in chemical fibrinolysis and thrombolysis, pp.3:191-209,
RL Raven Press, New York (1978).
RN [17]
RP SEQUENCE OF 483-604.
RX MEDLINE=76043692; PubMed=126663;
RA Wiman B., Wallen P.;
RT "Amino-acid sequence of the cyanogen-bromide fragment from human
RT plasminogen that forms the linkage between the plasmin chains.";
RL Eur. J. Biochem. 58:539-547(1975).
RN [18]
RP SEQUENCE OF 581-810.
RX MEDLINE=77225245; PubMed=142009;
RA Wiman B.;
RT "Primary structure of the B-chain of human plasmin.";
RL Eur. J. Biochem. 76:129-137(1977).
RN [19]
RP ACTIVE SITE.
RX MEDLINE=73149248; PubMed=4694729;
RA Robbins K.C., Bernabe P., Arrazon L., Summaria L.;
RT "Type primary structure of human plasminogen. II. The histidine loop
RT of human plasmin: light (B) chain active center histidine sequence.";
RL J. Biol. Chem. 248:1631-1633(1973).
RN [10]
RP ACTIVE SITE.
RX MEDLINE=69234739; PubMed=4240117;
RA Groskopf W.R., Summaria L., Robbins K.C.;
RT "Studies on the active center of human plasmin. Partial amino acid
RT sequence of a peptide containing the active center serine residue.";
RL J. Biol. Chem. 244:3590-3597(1969).
RN [11]
RP OMEGA-AMINOCARBOXYLIC ACID-BINDING SITES.
RX MEDLINE=8221905; PubMed=6919539;
RA Trexler M., Vail Z., Patchy L.;

RT "Structure of the omega-aminocarboxylic acid-binding sites of human
RT plasminogen. Arginine 70 and aspartic acid 56 are essential for
RT binding of ligand by kringle 4.";
RL J. Biol. Chem. 257:7401-7406(1982).
RN [12]
RP FIBRIN AND OMEGA-AMINOCARBOXYLIC ACID BINDING SITES.
RX MEDLINE=85054794; PubMed=6094526;
RA Vail Z., Patchy L.;
RT "The fibrin-binding site of human plasminogen. Arginines 32 and 34
RT are essential for fibrin affinity of the kringle 1 domain.";
RL J. Biol. Chem. 259:13690-13694(1984).
RN [13]
RP PHOSPHORYLATION SITE SER-597.
RX MEDLINE=97345939; PubMed=9201958;
RA Wang H., Proctor M., Brethauer R.K., Castellino F.J.;
RT "Serine-578 is a major phosphorylation locus in human plasma
RT plasminogen.";
RL Biochemistry 36:8100-8106(1997).
RN [14]
RP CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=8818329; PubMed=3356193;
RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamberling J.P.,
RT Gerwig G.J., van Halbeek H., Vliegenthart J.F.;
RT "The N- and O-linked carbohydrate chains of human, bovine and porcine
RT plasminogen. Species specificity in relation to stialylation and
RT fucoylation patterns.";
RL Eur. J. Biochem. 173:57-63(1988).
RN [15]
RP CARBOHYDRATE-LINKAGE SITE SER-268.
RX MEDLINE=97207306; PubMed=9054441;
RA Piate-Shepherd S.R., Stevens R.D., Andon N.L., Enghild J.J.,
RT Pizzo S.V.;
RT "Evidence for a novel O-linked sialylated trisaccharide on Ser-248 of
RT human plasminogen 2.";
RL J. Biol. Chem. 272:7408-7411(1997).
RN [16]
RP CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.
RX MEDLINE=95042728; PubMed=7525077;
RA O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,
RA Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
RT "Angiostatin: a novel angiogenesis inhibitor that mediates the
RT suppression of metastases by a Lewis lung carcinoma.";
RL Cell 79:315-328(1994).
RN [17]
RP CHARACTERIZATION OF ANGIOSTATIN.
RX MEDLINE=9728710; PubMed=9102221;
RA Sim B.K., O'Reilly M.S., Liang H., Fortier A.H., He W., Madsen J.W.,
RA Lapcevic R., Nacy C.A.;
RT "A recombinant human angiostatin protein inhibits experimental primary
RT Cancer Res. 57:1329-1334(1997).
RN [18]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 374-461.
RX MEDLINE=92031502; PubMed=1657148;
RA Mulichak A.M., Tullinsky A., Ravichandran K.G.;
RT "Crystal and molecular structure of human plasminogen kringle 4
RT refined at 1.9-A resolution.";
RL Biochemistry 30:10576-10586(1991).
RN [19]
RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 374-461.
RX MEDLINE=92031503; PubMed=1657149;
RA Wu T.-P., Padmanabhan K., Tullinsky A., Mulichak A.M.;
RT "The refined structure of the epsilon-aminocaproic acid complex of
RT human plasminogen kringle 4.";
RL Biochemistry 30:10589-10594(1991).
RN [20]
RP X-RAY CRYSTALLOGRAPHY (1.67 ANGSTROMS) OF 376-454.
RA Sec B., Yamano A., Whitlow M., Teeter M.M.;
RT "Structure of human plasminogen kringle 4 at 1.66 Angstrom and 277 K.
RT A possible structural role of disordered residues.";
RL Acta Crystallogr. D 53:169-178(1997).
RN [21]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 102-181.

RA MEDLINE=96180681; PubMed=8611560;
RA Mathews I.I., Vanderhoff-Hanaver P., Castellino F.J., Tulinsky A.;
RT "Crystal structures of the recombinant kringle 1 domain of human
RT plasminogen in complexes with the ligands epsilon-aminocaproic acid
RT and trans-4-(aminomethyl)cyclohexane-1-carboxylic Acid.";
RL Biochemistry 35:2567-2576(1996).
RN [22]
RP X-RAY CRYSTALLOGRAPHY (1.66 ANGSTROMS) OF 480-563.
RX MEDLINE=98198034; PubMed=9521645;
RA Chang Y., Mochalkin I., McCance S.G., Cheng B., Tulinsky A.,
RA Castellino F.J.;
RT "Structure and ligand binding determinants of the recombinant kringle
RT 5 domain of human plasminogen.";
RL Biochemistry 37:3258-3271(1998).
RN [23]
RP STRUCTURE BY NMR OF 96-184.
RX MEDLINE=94237157; PubMed=8181475;
RA Rejzance M.R., Llinas M.;
RT "1H-NMR assignments and secondary structure of human plasminogen
RT kringle 1.";
RL Eur. J. Biochem. 221:927-937(1994).
RN [24]
RP STRUCTURE BY NMR OF 96-184.
RX MEDLINE=94237158; PubMed=8181476;
RA Rejzance M.R., Llinas M.;
RT "Solution structure of the epsilon-aminohexanoic acid complex of
RT human plasminogen kringle 1.";
RL Eur. J. Biochem. 221:939-949(1994).
RN [25]
RP STRUCTURE BY NMR OF 183-354.
RX MEDLINE=96194156; PubMed=8652577;
RA Soehnlel S., Hu C.-K., Marti D., Affolter M., Schaller J., Llinas M.,
RA Ricci E.B.;
RT "Recombinant gene expression and 1H NMR characteristics of the
RT kringle (2 + 3) supermodule: spectroscopic/functional individuality
RT of plasminogen kringle domains.";
RL Biochemistry 35:2357-2364(1996).
RN [26]
RP STRUCTURE BY NMR OF 374-461.
RX MEDLINE=90219023; PubMed=2157850;
RA Atkinson R.A., Williams R.J.P.;
RT "Solution structure of the kringle 4 domain from human plasminogen by
RT 1H nuclear magnetic resonance spectroscopy and distance geometry.";
RL J. Mol. Biol. 212:541-552(1990).
RN [27]
RP VARIANTS PHE-374 AND THR-620.

Query Match 11.0%; Score 142; DB 1; Length 810;
Best Local Similarity 37.2%; Pred. No. 0.00016;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

QY 4 CFWDNGHYREDQSPAPGLRCUNWMDA-----QSGLASAPVSGAGNHSYCRNPDPDPG 58
DB 103 CTTGNGKRYRGTMTKTKGTCCKMSSTSHRRPRFPAHPEGL-ENNYCRNPDPDPG 161
QY 59 PWCYVSGEAGVPEKR--PCEDLRCPPE 82
DB 162 PWCYTTD---PEKRYDYCILCECE 183

RESULT 8
PLAN RAT STANDARD; PRT; 169 AA.
ID PLAN RAT
AC 001177;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Plasminogen (BC 3.4.21.7) (Fragment).
GN Plg.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=91250378; PubMed=1645711;
RA Karala J.J., Mäkelä S.P.;
RT Identification of the rat Heymann nephritis autoantigen (GP330) as a
RT receptor site for plasminogen.";
RL J. Biol. Chem. 266:10825-10829(1991).
CC -1- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC GRANULOSA FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC AS C1 AND C5. IT CLEAVES FIBRIN FIBRONECTIN, THROMBOSPONDIN,
CC LAMININ AND VON WILLEBRAND FACTOR.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -1- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC -1- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL, M62832, AAA41884.1, -
DR F1R; A40522; A40522.
DR HSSE; P00747; IPMK.
DR MEROP6; S01.233; -.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser protease_Try.
DR Pfam: PF00051; Kringle; 2.
DR ProDom: PD000395; Kringle; 2.
DR SMART: SMO0130; KR; 1.
DR PROSITE, PS00021; KRINGLE_1; 1.
DR PROSITE, PS00070; KRINGLE_2; 2.
DR PROSITE, PS50240; TRYPsin_DOM; PARTIAL.
DR PROSITE, PS50134; TRYPsin_HIS; PARTIAL.
DR PROSITE, PS00135; TRYPsin_SER; PARTIAL.
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Repeat.
FT NON TER 1
FT DOMAIN 1 10 KRINGLE 3 (BY SIMILARITY).
FT DOMAIN 34 112 KRINGLE 4 (BY SIMILARITY).
FT DOMAIN 139 >169 KRINGLE 5 (BY SIMILARITY).
FT DISULFD 34 112 BY SIMILARITY.
FT DISULFD 55 95 BY SIMILARITY.
FT DISULFD 83 107 BY SIMILARITY.
FT NON TER 169
SQ SEQUENCE 169 AA; 18401 MW; 77A54214C9D010C CRC64;

Query Match 10.7%; Score 138; DB 1; Length 169;
Best Local Similarity 31.8%; Pred. No. 6.3e-05;
Matches 35; Conservative 13; Mismatches 48; Indels 14; Gaps 5;

QY 4 CFWDNGHYREDQSPAPGLRCUNWMDA-----QSGLASAPVSGAGNHSYCRNPDPDPG 58
DB 34 CYQNGKXSYRGTSTTWYTKKCSQSWSMTPHSHSKTPANFPGSL-EMNYCRNPDPDPG 92
QY 59 PWCYVSGEAGVPEKR--PCEDLRCPPTSQALPFTTEIGQASGPGADP 106
DB 93 PWCYTTD---PSVMEYCNLKRCSEFGGV--AESAIIVQVPSAPGTSE 136

```

RESULT 9
UROC_PAPCY STANDARD; PRT; 433 AA.
ID UROC_PAPCY
AC P16227;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
GN (U-plasminogen activator).
OS Papio cynocephalus (Yellow baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Papio.
OX NCBI_TaxID=9556;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thoracic aorta;
RX MEDLINE=90287734; PubMed=2113276;
RA Au Y.P.T., Wang T.W., Clowes A.W.;
RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type
RT plasminogen activator."
RL Nucleic Acids Res. 18:3411-3411(1990).
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X51935; CAA36200.1; -.
DR PIR: S14687; UKBAY.
DR HSSP: P00749; 1LMM.
DR MEROPS: S01.231; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser protease_Try.
DR Pfam: PF00051; Kringle1; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR ProDom: PD000395; Kringle1; 1.
DR SMART: SM00181; EGF_1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; Tryp_Spc; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; FALSE NEG.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS50070; KRINGLE_2; 1.
DR PROSITE: PS50240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KM Plasminogen activation: Hydrolyase; Serine protease; Glycoprotein;
KM Kringle; EGF-like domain; Zymogen; Signal.
FT CHAIN 1 20 POTENTIAL.
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 176 CHAIN A (BY SIMILARITY).
FT CHAIN 155 176 SHORT A CHAIN (A1) (BY SIMILARITY).
FT CHAIN 178 433 CHAIN B (BY SIMILARITY).
FT CHAIN 26 62 EGF-LIKE.
FT DOMAIN 69 150 KRINGLE.

```

```

FT DOMAIN 151 177 CONNECTING PEPTIDE.
FT DISULFID 178 433 SERINE PROTEASE.
FT DISULFID 30 38 BY SIMILARITY.
FT DISULFID 32 50 BY SIMILARITY.
FT DISULFID 52 61 BY SIMILARITY.
FT DISULFID 167 298 INTERCHAIN (BY SIMILARITY).
FT DISULFID 208 224 BY SIMILARITY.
FT DISULFID 216 287 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 223 223 CHARGE RELAY SYSTEM.
FT ACT_SITE 274 274 CHARGE RELAY SYSTEM.
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM.
FT CARBOHYD 324 324 N-LINKED (GLCNAC...) (BY SIMILARITY).
SQ SEQUENCE 433 AA; 48595 MW; 816D22DFD0C8792 CRC64;

Query Match 10.7%; Score 198; DB 1; Length 433;
Best Local Similarity 32.7%; Pred. No. 0.00018;
Matches 32; Conservative 14; Mismatches 36; Indels 16; Gaps 4;

QY 4 CFWNGHLYREDQTSPPGLRCLNMLDA-----QSGLASPVSGAGHSYCRNDEDP 56
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 69 CYEGNGHYRKGKASTDWTGRSCLAAMSATVLDQTYHARSDALDGLGKKNYGRND-NR 127
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 57 RGPWCYSGEAGVPEK-----RPCEDLRCPEPITSQAL 88
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 128 RRPWCYV--QVGLKQRYQECVHNCADGKKFSSPEEL 163
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
UROC_PIG STANDARD; PRT; 442 AA.
ID UROC_PIG
AC P04185;
DT 20-MAR-1987 (Rel. 04, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
GN (U-plasminogen activator).
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
OX NCBI_TaxID=9623;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=85087954; PubMed=6096832;
RA Nagamine Y., Pearson D., Altus M.S., Reich E.;
RL "cDNA and gene nucleotide sequence of porcine plasminogen activator."
RL Nucleic Acids Res. 12:9525-9541 (1984).
RN [2]
RP REVISION TO 241.
RA Nagamine Y.
RU Submitted (DEC-1986) to the PIR data bank.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X01648; CAA25806.1; -.
DR EMBL: X02724; CAA26511.1; -.
DR PIR: A00932; UKPG.
DR HSSP: P00749; 1KDU.

```


DR MEROPS; S01.231; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYP_SPE; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00021; KRINGE_1; 1.
DR PROSITE; PS00070; KRINGE_2; 1.
DR PROSITE; PS00240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
DR Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20
FT CHAIN 21 442 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 188 CHAIN A (BY SIMILARITY).
FT CHAIN 190 442 CHAIN B (BY SIMILARITY).
FT DOMAIN 29 65 EGF-LIKE.
FT DOMAIN 72 153 KRINGE.
FT DOMAIN 154 189 CONNECTING PEPTIDE.
FT DOMAIN 190 442 SERINE PROTEASE.
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .).
FT DISULFID 33 41 BY SIMILARITY.
FT DISULFID 35 53 BY SIMILARITY.
FT DISULFID 55 64 BY SIMILARITY.
FT DISULFID 179 310 INTERCHAIN (BY SIMILARITY).
FT DISULFID 220 236 BY SIMILARITY.
FT DISULFID 228 299 BY SIMILARITY.
FT DISULFID 324 393 BY SIMILARITY.
FT DISULFID 356 372 BY SIMILARITY.
FT DISULFID 383 411 BY SIMILARITY.
FT ACT_SITE 235 235 CHARGE RELAY SYSTEM.
FT ACT_SITE 286 286 CHARGE RELAY SYSTEM.
FT ACT_SITE 387 387 CHARGE RELAY SYSTEM.
FT ACT_SITE 241 241 O -> H (IN REF. 1; CAA25806).
FT CONFLICT 242 242 O -> H (IN REF. 1; CAA25811).
FT CONFLICT 288 288 A -> GS (IN REF. 1; CAA25806).
SQ SEQUENCE 442 AA; 49116 MM; E832FCFE501321EB CRC64;

Query Match 10.5%; Score 135.5; DB 1; Length 442;
Best Local Similarity 36.9%; Pred. No. 0.0003;
Matches 38; Conservative 7; Mismatches 37; Indels 21; Gaps 5;

QY 4 CFWNHGHLVREDQSPAPGRCLNWLDAQSG-----ASAPVS-----GAGNHSYCRNPEDP 56
DB 72 CFEENGHSYRKANTNTNGRRCPLPMSATVLTNTYAHKPDALQGLKHKYCNPPD-NO 130
QY 57 RGPWCYVS-----GEAGVP-----EKRPCEDLRCPTTSQ 86
DB 131 RRPWCYVGVGLKQLVQECWPNCSGSGSHRPAYDGKPPSTBE 173

RESULT 11
URTB DESRO
ID _URTB_DESRO STANDARD; PRT; 431 AA.
AC P98121.
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator beta precursor (EC 3.4.21.68) (DSPA
beta).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;

EN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "the plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Wilt W.,
RA Donner P.;
RT "plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity."
RL Ann. N.Y. Acad. Sci. 667:385-403(1992).
CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
CC AGENT.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: MONOMER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTRAINS 1 KRINGE DOMAIN.
CC -1- SIMILARITY: CONTRAINS 1 EGF-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M63989; AAA31594.1; -.
DR HSSP; P98119; 1A51.
DR MEROPS; S01.239; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser protease_Try.
DR Pfam; PF00089; EGF; 1.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYP_SPE; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00021; KRINGE_1; 1.
DR PROSITE; PS00070; KRINGE_2; 1.
DR PROSITE; PS00240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Multigene family.
FT SIGNAL 1 36
FT CHAIN 37 431 SALIVARY PLASMINOGEN ACTIVATOR BETA.
FT DOMAIN 37 75 EGF-LIKE.
FT DOMAIN 82 163 KRINGE.
FT DOMAIN 179 431 SERINE PROTEASE.
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 275 275 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 382 382 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 41 52 BY SIMILARITY.
FT DISULFID 46 63 BY SIMILARITY.
FT DISULFID 65 74 BY SIMILARITY.
FT DISULFID 82 163 BY SIMILARITY.


```
FT CONFLICT 403 403 N -> K (IN REF. 2).
FT CONFLICT 417 417 Y -> H (IN REF. 2).
FT CONFLICT 435 435 M -> R (IN REF. 2).
SQ SEQUENCE 477 AA; 53719 MW; 17486555C0E5077C CRC64;

Query Match 10.4%; Score 134; DB 1; Length 477;
Best Local Similarity 38.6%; Pred. No. 0.00044;
Matches 27; Conservative 9; Mismatches 20; Indels 14; Gaps 3;

OY 4 CPWDGHLHYREDQSPAPGLRGLTLMNLDLQSGI-----ASAPVSGAGNHSYCRNPD 53
Db 128 CKXDGVTYRGWSTSGSQAQCTNW---NSNLTTRTYNGRRSDAITLGLGNHNTCRNPD 184
OY 54 EDPGRPCVY 63
Db 185 NNSK-PWCYV 193

RESULT 13
UROK_HUMAN STANDARD; PRT; 431 AA.
ID _UROK_HUMAN 015844; Q16618; Q969W6;
AC 21-JUL-1986 (Rel. 01, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)
DE (U-plasminogen activator).
GN PLAV.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=85215647; PubMed=2987867;
RA Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Biasi F.;
RT "The human urokinase-plasminogen activator gene and its promoter.";
RL Nucleic Acids Res. 13:2759-2771(1985).
RN 12
RP SEQUENCE FROM N.A.
RA Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,
RA Steffens G.J., Heynaker H.U.;
RT "Cloning and expression of the gene for pro-urokinase in Escherichia
RT coli.";
RL Biotechnology 3:923-929(1985).
RN 13
RP SEQUENCE FROM N.A.
RX MEDLINE=86056954; PubMed=2415429;
RA Nagai M., Hiramatsu R., Kaneda T., Hayasuke N., Arimura H.,
RA Nishida M., Suyama T.;
RT "Molecular cloning of cDNA coding for human preprourokinase.";
RL Gene 36:183-188(1985).
RN 14
RP SEQUENCE FROM N.A.
RX MEDLINE=85203359; PubMed=3888571;
RA Jacobs P., Cravador A., Lortiau R., Brockly F., Coliau B., Chuchana P.,
RA van Elsen A., Herzog A., Bollen A.;
RT "Molecular cloning, sequencing, and expression in Escherichia coli of
RT human preprourokinase cDNA.";
RL DNA 4:139-146(1985).
RN 15
RP SEQUENCE FROM N.A.
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA Nickerson D.A.;
RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN 16
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strusberg R.;
RN 17
RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RX MEDLINE=84272706; PubMed=6589620;

RA Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Biasi F.;
RT "Identification and primary sequence of an unspliced human urokinase
RT poly(A)+ RNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).
RN 18
RP SEQUENCE OF 21-177.
RX MEDLINE=83055084; PubMed=6754569;
RA Gunzier W.A., Steffens G.J., Oetting F., Kim S.-M.A., Frankus E.,
RA Flohe L.;
RT "The primary structure of high molecular mass urokinase from human
RT urine. The complete amino acid sequence of the A chain.";
RL Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).
RN 19
RP SEQUENCE OF 156-176 AND 179-224.
RX MEDLINE=83003608; PubMed=6749491;
RA Schaller J., Nick H., Rickl E.E., Gilleesen D., Lergier W.,
RA Studer R.O.;
RT "Human low-molecular-weight urinary urokinase. Partial
RT characterization and preliminary sequence data of the two polypeptide
RT chains.";
RL Eur. J. Biochem. 125:251-257(1982).
RN 110
RP SEQUENCE OF 158-410.
RX MEDLINE=83055099; PubMed=6754572;
RA Steffens G.J., Gunzier W.A., Oetting F., Frankus E., Flohe L.;
RT "The complete amino acid sequence of low molecular mass urokinase
RT from human urine.";
RL Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).
RN 111
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=96000858; PubMed=8591045;
RA Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,
RA Dobson C.M., Stuart D.I., Jones E.Y.;
RT "The crystal structure of the catalytic domain of human
RT urokinase-type plasminogen activator.";
RL Structure 3:681-691(1995).
RN 112
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.
RX MEDLINE=2026327; PubMed=10805774;
RA Speer S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G.,
RA Bode W., Magdolen V., Huber R., Moroder L.;
RT "(4-aminomethyl)phenylguanidine derivatives as nonpeptidic highly
RT selective inhibitors of human urokinase.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).
RN 113
RP STRUCTURE BY NMR.
RX MEDLINE=89127526; PubMed=2536903;
RA Oswald R.E., Bogusky M.U., Bamberger M., Smith R.A.G., Dobson C.M.;
RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-
RT dimensional NMR.";
RL Nature 337:579-582(1989).
RN 114
RP STRUCTURE BY NMR OF 67-155.
RX MEDLINE=93003110; PubMed=13727118;
RA Li X., Smith R.A.G., Dobson C.M.;
RT "Sequential 1H NMR assignments and secondary structure of the kringle
RT domain from urokinase.";
RL Biochemistry 31:9562-9571(1992).
RN 115
RP STRUCTURE BY NMR OF 67-155.
RX MEDLINE=94149701; PubMed=8107091;
RA Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.;
RT "Solution structure of the kringle domain from urokinase-type
RT plasminogen activator.";
RL J. Mol. Biol. 235:1548-1559(1994).
RN 116
RP VARIANT LEU-141.
RX MEDLINE=96186279; PubMed=8652631;
RA Yoshimoto M., Ushiyama Y., Sakai M., Tamaki S., Hara H., Takehashi K.,
RA Sawasaki Y., Hanada K.;
RT "Characterization of single chain urokinase-type plasminogen
RT activator with a novel amino-acid substitution in the kringle
RT structure.";
```



```

DR PRIN7S; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AB; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS50070; KRINGLE_2; 4.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR KRingle; Glycoprotein; Serine protease homolog; Repeat; Signal.
KW SIGNAL.
FT CHAIN 1 18
FT CHAIN 19 716
FT CHAIN 109 109
FT CHAIN 110 186
FT CHAIN 191 268
FT CHAIN 292 370
FT CHAIN 379 457
FT CHAIN 489 716
FT CHAIN 56 78
FT CHAIN 60 66
FT CHAIN 110 186
FT CHAIN 131 169
FT CHAIN 157 181
FT CHAIN 191 268
FT CHAIN 194 333
FT CHAIN 212 251
FT CHAIN 240 263
FT CHAIN 292 370
FT CHAIN 313 352
FT CHAIN 341 364
FT CHAIN 379 457
FT CHAIN 400 440
FT CHAIN 428 452
FT CHAIN 472 593
FT CHAIN 512 528
FT CHAIN 607 672
FT CHAIN 637 651
FT CHAIN 662 690
FT CHAIN 72 72
FT CHAIN 173 173
FT CHAIN 305 305
FT CHAIN 620 620
FT CHAIN 19 19
SQ SEQUENCE 716 AA; 80588 MW; BBCE02EF85213ACC CRC64;

Query Match 10.1%; Score 130.5; DB 1; Length 716;
Best Local Similarity 30.7%; Pred. No. 0.0014;
Matches 35; Conservative 7; Mismatches 37; Indels 35; Gaps 5;

OY 4 CFWDNGHLYREDQTSAPGLRCLNW---LDAQSGLASAFVSGAGNHSYCRNPDEDPGRFW 60
DB 110 CIMDNGVSRGTVARTAGLPCQAMRRFPNDHKYTPPKNGL-BENFCRNDGDPGRFW 168

OY 61 CYVS-----GEAGVPEK-RPCE--DIRCET 83
DB 169 CYTTHSVRFQSGIKTCREAVCVLCNGEDYGEVDVTEGREGCRWDLQHPHS 222

RESULT 15
PLMN_ERIEU STANDARD; PRT; 810 AA.
AC 029485;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Plasmidogen precursor (EC 3.4.21.7).
GN PLG.
OS Brinaceus europaeus (Western European hedgehog).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Mammalia; Eutheria; Insectivora; Erinaceidae; Erinaceinae; Erinaceus.
OX NCBI_TaxID=9365;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;

```

```

RX MEDLINE=96025778; PubMed=7592597;
RA Lawn R.M., Boommark N.W., Schwartz K., Lindahl G.E., Wade D.P.,
RA Byrne C.D., Fong K.U., Meer K., Patchy L.;
RT "The recurring evolution of lipoprotein(a). Insights from cloning of
RL hedgehog apolipoprotein(a).";
RN J. Biol. Chem. 270:24004-24009(1995).
RP REVISIONS.
RA Lawn R.M.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC GRAFIAN FOLLICLE. IT ACTIVATES THE UKRINASE-TYPE PLASMINOGEN
CC ACTIVATOR. COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC AS C1 AND C5, IT CLEAVES FIBRIN, FIBRINECTIN, THROMBOSPONDIN,
CC LAMININ AND VON WILLEBRAND FACTOR.
CC -1- CATALYTIC ACTIVITY: preferential cleavage: Lys-|-Xaa -> Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -1- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC ACTIVATORS. BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC -1- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U3171; AAC48717.1; -.
DR HSSP; P00747; 1PMK.
DR MEROPS; S01.233; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PR00024; PAN; 1.
DR Pfam; PR00051; kringle; 5.
DR Pfam; PF00089; chymotrypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 5.
DR SMART; SM00473; PAN_AB; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 5.
DR PROSITE; PS50070; KRINGLE_2; 5.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR HydroLase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
KW Signal.
FT SIGNAL 1 19
FT CHAIN 20 810
FT CHAIN 20 582
FT CHAIN 583 810
FT CHAIN 583 810
FT CHAIN 103 810
FT CHAIN 181 181
FT CHAIN 185 262
FT CHAIN 275 352
FT CHAIN 379 456
FT CHAIN 482 561
FT ACT_SITE 622 622
CC BY SIMILARITY.
CC PLASMINOGEN.
CC PLASMIN HEAVY CHAIN A (BY SIMILARITY).
CC PLASMIN LIGHT CHAIN B (BY SIMILARITY).
CC SERINE PROTEASE.
CC KRINGLE 1.
CC KRINGLE 2.
CC KRINGLE 3.
CC KRINGLE 4.
CC KRINGLE 5.
CC CHARGE RELAY SYSTEM.

```

FT ACT SITE 665 665 CHARGE RELAY SYSTEM.
 FT ACT SITE 760 760 CHARGE RELAY SYSTEM.
 FT CAREOHD 339 339 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 810 AA; 90902 MM; BE75780946017A16 CRC64;

Query Match 10.0%; Score 129; DB 1; Length 810;
 Best Local Similarity 34.1%; Pred. No. 0.0021;
 Matches 31; Conservative 7; Mismatches 31; Indels 22; Gaps 4;

Qy 4 CFWDNGHLYREDQTSBPGLRCLNWLDAOSGLASAP-----VSGAGNHSYCRNPD 53
 | | | | | : | | | | | : | | | | |
 Db 103 CRYGNGKYYRGTVSKTKGLTCOKW-----SAETPHKRFSPDENPSEGLDQNYCRNPD 156
 | | | | | : | | | | | : | | | | |
 Qy 54 EDPGPGWCYVSGEAGVPEKR--PCEDLRQPE 82
 | | | | | : | | | | | : | | | | |
 Db 157 NDPKGPWCYTMD---PEVRYEYCEIIQCED 183
 | | | | | : | | | | | : | | | | |

Search completed: April 7, 2003, 09:16:56
 Job time : 13.0686 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:25 ; Search time 64.9023 Seconds

(without alignments)
768.284 Million cell updates/sec

Title: US-10-057-951-2_COPY_22_263

Sequence: 1 SGGCFWMDNGHLTYREDQTSFA.....PVDQEGSTPLMGAGATPGCA 242

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1289	100.0	263	4 Q96FE7	Q96FE7 homo sapien
2	1286	99.8	263	4 O00318	O00318 homo sapien
3	167	13.0	562	6 O8S023	O8S023 sus scrofa
4	154.5	12.0	516	4 O9BU99	O9BU99 homo sapien
5	150	11.6	653	11 O8VCS4	O8VCS4 mus musculu
6	147.5	11.4	559	11 O91VP2	O91VP2 mus musculu
7	142	11.0	810	4 Q15146	Q15146 homo sapien
8	136.5	10.6	420	13 O90504	O90504 epistretus
9	135	10.5	385	5 Q25101	Q25101 hermania m
10	135	10.5	812	11 O9R0W3	O9R0W3 ratus norv
11	134.5	10.4	704	13 O90865	O90865 gallus gall
12	133.5	10.4	716	11 P70521	P70521 ratus norv
13	131.5	10.2	313	13 O9PU78	O9PU78 crocodylus
14	131	10.2	154	4 O96S88	O96S88 homo sapien
15	131	10.2	608	13 O9PTW7	O9PTW7 struthio ca
16	130.5	10.1	716	11 Q91XG8	Q91XG8 mus musculu

17	130.5	10.1	717	13 P70006	P70006 xenopus lae
18	129	10.0	616	6 O97507	O97507 sus scrofa
19	128.5	10.0	806	6 O18783	O18783 macropus eu
20	127	9.9	157	6 O9TV48	O9TV48 bos taurus
21	127	9.9	395	4 O9BZW1	O9BZW1 homo sapien
22	126	9.8	716	13 O91691	O91691 xenopus lae
23	123	9.5	728	6 O9BH09	O9BH09 felle silve
24	119.5	9.3	399	4 O96GL8	O96GL8 homo sapien
25	119.5	9.3	420	4 O9BTP9	O9BTP9 homo sapien
26	119.5	9.3	560	4 O14520	O14520 homo sapien
27	119	9.2	334	6 O46507	O46507 papio hamad
28	118	9.2	710	13 O91402	O91402 xenopus. he
29	117.5	9.1	812	11 O91WJ5	O91WJ5 mus musculu
30	114.5	8.9	113	4 O9UR17	O9UR17 homo sapien
31	114	8.8	113	4 O9UR5	O9UR5 homo sapien
32	114	8.8	202	13 O90675	O90675 gallus gall
33	114	8.8	607	13 O91001	O91001 gallus gall
34	113.5	8.8	452	13 O90Y90	O90Y90 gallus gall
35	112.5	8.7	359	6 O8WNR1	O8WNR1 canis fami
36	112	8.7	567	4 O13208	O13208 homo sapien
37	112	8.7	594	5 P91823	P91823 caenorbadi
38	112	8.7	1145	5 O9BK18	O9BK18 alysia cal
39	111.5	8.7	397	11 O35727	O35727 mus musculu
40	111	8.6	726	13 O90978	O90978 gallus gall
41	110.5	8.6	685	5 Q24488	Q24488 drosophila
42	110	8.5	2358	16 O91LV8	O91LV8 streptomyce
43	108	8.4	215	13 O42341	O42341 gallus gall
44	106	8.2	378	13 O90WP0	O90WP0 trachemys s
45	105.5	8.2	709	13 Q90ZM6	Q90ZM6 brachydario

ALIGNMENTS

RESULT 1
Q96FE7 PRELIMINARY; PRT; 263 AA.
ID Q96FE7
AC Q96FE7;
DC 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Unknown (protein for MGC:17330).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC011049; AAH11049.1; -;
DR InterPro: IPR00001; Kringle.
DR Pfam: PF00051; Kringle; 1.
DR ProDom: PD000395; Kringle; 1.
DR PROSITE: PS00021; Kringle_1; UNKNOWN_1.
DR PROSITE: PS50070; Kringle_2; 1.
SQ
SEQUENCE 263 AA; 28234 MW; 197C3EE888FA242 CRC64;

Query Match 100.0%; Score 1289; DB 4; Length 263;
Best Local Similarity 100.0%; Pred. No. 9.5e-109;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	SGGCFWMDNGHLTYREDQTSFA	GCRCNMTDAOSGLASAPVSGAGNHSYCRNPDEDRGW	60
DB	22	SGGCFWMDNGHLTYREDQTSFA	GCRCNMTDAOSGLASAPVSGAGNHSYCRNPDEDRGW	81
QY	61	CYVSGEAGVEKRCPCDILRCPTTSQALPAFTTEIOEASGCADEVQVAFANALPARS		120
DB	82	CYVSGEAGVEKRCPCDILRCPTTSQALPAFTTEIOEASGCADEVQVAFANALPARS		141
QY	121	EAANAQVPIGISQRYVMSKREKDLGTGLGYVIGITMVIIIAIGAGIILGYSYKRGKDLK		180

```

Db 142 EAAAVPVIGISORVRNNSKEKDLGLGYLGITMMVIIAIGAGIILGYSYKRGDLK 201
QY 181 EOHDOKVCEREMORTITPLSAFTNPCTCEIVDEKTVVHTSQTPVDPOEGSTPLMGQAGTP 240
Db 202 EOHDOKVCEREMORTITPLSAFTNPCTCEIVDEKTVVHTSQTPVDPOEGSTPLMGQAGTP 261
QY 241 GA 242
Db 262 GA 263

RESULT 2
ID 000318 PRELIMINARY; PRT; 263 AA.
AC 000318;
DT 01-JUN-1997 (TReMBLrel. 04, Created)
DT 01-JUN-1997 (TReMBLrel. 04, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE WUSGC.DJ515N1.2 protein.
GN WUSGC.DJ515N1.2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OK NCBI_Taxid=9606;
RN SEQUENCE FROM N.A.
RP Du Z., Scheet P., Harper M.;
RT "The sequence of H. sapiens PAC clone RP3-515N1."
RL Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AC002073; AAB54054.1; -.
DR HSSP; P00749; 1XDU.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE; FALSE_NEG.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
SQ SEQUENCE 263 AA; 28248 MW; 197C3EEBE854A242 CRC64;

Query Match 99.8%; Score 1286; DB 4; Length 263;
Best Local Similarity 99.6%; Pred. No. 1.8e-108;
Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGGCFMDNGHLVREDQTSAPGLRCINWIDAQSLASAPVSGAGNHSYCRNDEDEPRGPW 60
Db 22 SGGCFMDNGHLVREDQTSAPGLRCINWIDAQSLASAPVSGAGNHSYCRNDEDEPRGPW 81
QY 61 CYYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASEGPGADEVQYFAPANALPARS 120
Db 82 CYYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASEGPGADEVQYFAPANALPARS 141
QY 121 EAAAVPVIGISORVRNNSKEKDLGLGYLGITMMVIIAIGAGIILGYSYKRGDLK 180
Db 142 EAAAVPVIGISORVRNNSKEKDLGLGYLGITMMVIIAIGAGIILGYSYKRGDLK 201
QY 181 EOHDOKVCEREMORTITPLSAFTNPCTCEIVDEKTVVHTSQTPVDPOEGSTPLMGQAGTP 240
Db 202 EOHDOKVCEREMORTITPLSAFTNPCTCEIVDEKTVVHTSQTPVDPOEGSTPLMGQAGTP 261
QY 241 GA 242
Db 262 GA 263

RESULT 3
ID 085023 PRELIMINARY; PRT; 562 AA.
AC 085023;

```

```

DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE T-plasminogen activator.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OK NCBI_Taxid=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ENAMEL ORGAN;
RA Ding Y., Xue J., Bartlett J.D.;
RT "T-plasminogen activator in tooth tissues."
RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF364605; AAM00297.1; -.
SQ SEQUENCE 562 AA; 63668 MW; F9B6B4C77CB101E8 CRC64;

Query Match 13.0%; Score 167; DB 6; Length 562;
Best Local Similarity 36.6%; Pred. No. 1e-06;
Matches 41; Conservative 12; Mismatches 33; Indels 26; Gaps 6;

QY 4 CFMDNGHLVREDQTSAPGLRCINWIDAQSLASAPVSGAGNHSYCRNDEDEPRGPW 54
Db 127 CYBDQGITRYGTWSTTESGAEQVNW--NTSGLASMPYNGRRPDAYKLGNGHNCRNPDK 184
QY 55 DPRGPMCVVSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASEGPGAD 105
Db 185 DSK-PWCYITFPAKTSYD-----FC-----STPACTKEKEBECTGTGGLD 222

RESULT 4
ID 09B099 PRELIMINARY; PRT; 516 AA.
AC 09B099;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Similar to plasminogen activator, tissue.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OK NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SKIN;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, ALSO KNOWN AS THE TRYPSIN FAMILY.
DR EMBL; BC002795; AAH02795.1; -.
DR HSSP; P00750; 1A5H.
DR MEROPS; S01.232; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00051; Kringle; 2.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00001; EGF-like; 1.
DR SMART; SM00001; EGF-like; 1.
DR SMART; SM000130; KR; 2.
DR SMART; SM00020; TRY_5PC; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01166; EGF_2; 1.
DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS0134; TRYPsin_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPsin_SER; 1.

```



```

DE Plasminogen precursor.
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Byrme M.J., Chapman C.G., Dodd I., Carey J.E., Lawrence G.M.P.,
RT Mitchell D., Robinson J.H.;
RT "Expression of recombinant human plasminogen and aglycoplasminogen in
RT HeLa cells.";
RT Fibrinolysis 0:0-0(1991).
CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: M74220; AAA36451.1; -.
DR HSP: P00747; 2PK4.
DR MEROPS: S01.213; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan_app.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00051; Kringle_5.
DR Pfam: PF00024; PAN; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle_5.
DR SMART: SM00130; KR; 5.
DR SMART: SM00473; PAN_Ap; 1.
DR PROSITE: PS00020; TRYP_SPC; 1.
DR PROSITE: PS00021; KRINGLE_1; 5.
DR PROSITE: PS50070; KRINGLE_2; 5.
DR PROSITE: PS50240; TRYPsin_DOM; 1.
DR PROSITE: PS00134; TRYPsin_HIS; 1.
DR PROSITE: PS00135; TRYPsin_SER; 1.
DR KX Hydroxylase; Serine protease; Signal.
FT SIGNAL 1 19
FT CHAIN 20 810 PLASMINOGEN.
SQ SEQUENCE 810 AA; 90555 MW; B05C7D4B0D020B3C CRC64;

Query Match 11.0%; Score 142; DB 4; Length 810;
Best Local Similarity 37.2%; Pred. No. 0.0003;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

QY 4 CFWDNGHLYREDQTSAPAGRLCLNWLDA-----QSGLASAPVSGAGNHSYCRNPDEDPDG 58
DB 103 CKTGNGKNYRGWMSKTKNGKITCQMSSTSPHRPRFSPATHPSEGL-ENNYCRNPNDNDPOG 161
QY 59 PWCYVSGEAGVPEKR--PCEDLRCP 82
DB 162 PWCYTTD---PEKKYDYCDILECEE 183

```

RESULT 8

```

ID Q90504 PRELIMINARY; PRT; 420 AA.
AC Q90504;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Thrombin.
OS Eptatretus scoulli (Pacific hagfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Hyperotreti; Myxiniiformes;
OC Myxiniidae; Eptatretinae; Eptatretus.
OX NCBI_TaxId=7765;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Banfield D.K., Macgillivray R.T.;
RA Banfield D.K., Macgillivray R.T.;
RA "Partial characterization of vertebrate prothrombin cDNAs

```

```

RT amplification and sequence analysis of the B chain of thrombin from
RT nine different species.";
RT Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
RN [2]
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=94223694; PubMed=7513365;
RA Banfield D.K., Irwin D.M., Walz D.A., Macgillivray R.T.;
RT "Evolution of prothrombin: isolation and characterization of the cDNAs
RT encoding chicken and hagfish prothrombin.";
RT J. Mol. Evol. 38:177-187(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Banfield D.K.;
CC Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.
CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: M81393; AAA21620.1; -.
DR HSP: P00734; 10V6.
DR MEROPS: S01.217; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00051; Kringle; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; TRYP_SPC; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS50070; KRINGLE_2; 1.
DR PROSITE: PS50240; TRYPsin_DOM; 1.
DR PROSITE: PS00134; TRYPsin_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPsin_SER; 1.
DR KX Hydroxylase; Serine protease.
SQ SEQUENCE 420 AA; 47888 MW; 64522AA21A57B67A CRC64;

```

```

Query Match 10.6%; Score 136.5; DB 13; Length 420;
Best Local Similarity 26.5%; Pred. No. 0.00041;
Matches 39; Conservative 16; Mismatches 77; Indels 15; Gaps 4;

```

```

QY 4 CFWDNGHLYREDQTSAPAGRLCLNWLDAQSGLASAPVSGAG-NHSYCRNPDEDPDGPNICY 62
DB 17 CYRREGDYRGDNLITWTKGKPLPWRGSYSNLPQFTTAGLTNSYCRNPDGUSRGWCY 76
QY 63 VSGEAGVP---EKRPCEDLRCPETTSQALPAFTTEIQEASBSPGADDEVAPANALPA 118
DB 77 TYGVEGTVDVDCOLNACE-----SGDIFVGTDEVQLSGRSEGAAEKTLFFNPKTFGN 129
QY 119 RSEAAAVPVIGISORVRNMSKKDL 145
DB 130 GEEGCKRPWFELQK---NDRSEDEL 153

```

RESULT 9

```

ID Q25101 PRELIMINARY; PRT; 385 AA.
AC Q25101;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Serine proteinase.
OS HMsERPp.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Fyuridae; Herdmania.
OX NCBI_TaxId=7733;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CURVATA;
RA Arnold J.M., Kennett C., Lavin M.F.;

```

RT "Transient expression of a novel serine protease in the ectoderm of
 RL the ascidian *Herdmania momus* during development.",
 CC Dev. Genes Evol. 206:455-463(1997).
 -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 TRYPSIN FAMILY.
 DR EMBL: U63517; AAB6650.1; -.
 DR HSSP: P00763; IDPO.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR001254; Ser. protease_Try.
 DR Pfam: PF00051; kringle_1.
 DR Pfam: PF00089; trypsin_1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00018; KRINGLE.
 DR ProDom: PD000395; Kringle_1.
 DR SMART: SM00130; KR_1.
 DR SMART: SM00020; TRYPSIN_SPC; 1.
 DR PROSITE: PS00021; KRINGLE_1; 1.
 DR PROSITE: PS50070; KRINGLE_2; 1.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 DR HydroLase: Serine protease.
 SQ SEQUENCE 385 AA; 42935 MW; BFBID05D5232E6A0 CRC64;

Query Match 10.5%; Score 135; DB 5; Length 385;
 Best Local Similarity 20.2%; Pred. No. 0.0005;
 Matches 65; Conservative 44; Mismatches 121; Indels 92; Gaps 12;

QY 4 CF-WDNGHLVREDQTSAPAGRLCNLWDAOS-----GLASAVSGAGNHSYCRNDDDEP 56
 DB 23 CFIENNESISQGAISRLTGETEQSW-DLQTPHKHKTYSNYSNGIAGNNVCRNDQDM 81
 QY 57 RGVWCYVSGE-----AGVPEKPCEDLRCPETT---SQALPAFTTEIOEASG 101
 DB 82 RGVWCYTNFEMWDYCDIPICSNPPVTLPSIEGKTEPELSDTKGDKLQSAKTN 141
 QY 102 P-----GADEVQ-----VFAPNALPARSEAAAVQPV 128
 DB 142 PLHIVGTTVTHGISIPWQYSLRLKRELRFHFCGSIINRWILFAHCIRPQPKYLA 201
 QY 129 IGISORFNMSEKEDLGLGYVL-----GITMVAIIAGAGIILGYSK 174
 DB 202 LGYVDALQVDFSEMK---VGFRLFNHKKYNPATENDITLTKMDTSSIAITFQGSVF 257
 QY 175 RGDLDKEQHDQVC-----EREMORITLPLSAFTNPTCEIIVEKTV---VYH 218
 DB 258 PPAKVPAAKSKTIIVSGWGDPTKGTQDVKLNVTLFVMSF--KLCKKLYSKVVGAAVPRK 315
 QY 219 TSGTPVDPQEGSTPLMGACTP 240
 DB 316 TSLCAAYKKGKSDSCQDSGGP 337

RESULT 10
 Q9ROW3 PRELIMINARY; PRT; 812 AA.
 AC Q9ROW3;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Plasmidogen protein precursor (EC 3.4.21.7).
 GN PLASMINOGEN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Bangert K.; Johnsen A.H.; Thorsen S.;
 RT "Rat plasmidogen: cDNA and gene structure."
 RT Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE=91250378; Pubmed=1645711;
 RA Kanalas J.J.; Makker S.P.;
 RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a
 RT receptor site for plasmidogen."
 RT J. Biol. Chem. 266:10825-10828 (1991).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 TRYPSIN FAMILY.
 DR EMBL: AJ242649; CAB46014.1; -.
 DR HSSP: P00747; LPMK.
 DR MEROPS: S01.233; -.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR003014; PAN.
 DR InterPro: IPR003609; Pan. app.
 DR InterPro: IPR001254; Ser. protease_Try.
 DR InterPro: IPR001400; Somatostatin.
 DR Pfam: PF00051; kringle_5.
 DR Pfam: PF00024; PAN_1.
 DR Pfam: PF00089; trypsin_1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00018; KRINGLE.
 DR ProDom: PD000395; Kringle_5.
 DR SMART: SM00130; KR; 4.
 DR SMART: SM00473; PAN_Ap; 1.
 DR SMART: SM00020; TRYPSIN_SPC; 1.
 DR PROSITE: PS00021; KRINGLE_1; 5.
 DR PROSITE: PS50070; KRINGLE_2; 5.
 DR PROSITE: PS50240; SOMATOTROPIN_2; UNKNOWN_1.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 DR HydroLase: Serine protease; Signal.
 FT SIGNAL 1
 FT CHAIN 1
 SQ SEQUENCE 812 AA; 90535 MW; 8C703C51410EBC9E CRC64;

Query Match 10.5%; Score 135; DB 11; Length 812;
 Best Local Similarity 30.9%; Pred. No. 0.0013;
 Matches 34; Conservative 14; Mismatches 48; Indels 14; Gaps 5;
 QY 4 CFWDNGHLVREDQTSAPAGRLCNLWDA-----QSLASAPVSGAGNHSYCRNDDDEP 58
 DB 376 CYQNGKSYKRTSTTTGKKCSWVSMTPHSKSTPAPNPDAGL-EMNYCRNDDQDG 434
 QY 59 FWCYSGEAGVPEKR--PCEDLRCPETTSQLPAFTTEIOEASGPGADE 106
 DB 435 FWCFTTD---PSVMEYCNLKRCSFTGGV--AESAIIVQVSAPOITSE 478

RESULT 11
 Q90865 PRELIMINARY; PRT; 704 AA.
 AC Q90865;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Hepatocyte growth factor-like/macrophage stimulating protein.
 GN HGF1/MSP.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE=96029010; Pubmed=7554499;
 RA Thery C.; Sharpe M.J.; Batley S.J.; Stern C.D.; Gherardi E.;
 RT "Expression of HGF/SF, HGF1/MSP and c-met suggests new functions
 RT during early chick development.";

```

DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
KW Hydrolyase; Serine protease; Signal.
FT SIGNAL 1
FT SEQUENCE 716 AA; 80733 MW; 06B7DF3BF56D921F CRC64;
SQ
Query Match 10.4%; Score 133.5; DB 11; Length 716;
Best Local Similarity 29.1%; Pred. No. 0.0015;
Matches 41; Conservative 9; Mismatches 48; Indels 43; Gaps 6;
QY 4 CFWNDGHHYREDQTSPPAGLRCLNM---LDAGSLASAPVSGAGNHSYCRNPDEDPRGPW 60
DB 110 CIMDNGASVSGTVAFTADGLPCQAKSRFRFPNDKHTPTPKNGL-EENFCRNPDDDPKGPW 168
QY 61 CVVS-----GAGVPEK-RPCE--DLRCPEF----- 83
DB 169 CYTTNRSVRFQSCGKIKSCREAVCVWCGNEDYRGVVDVTESGRECRWDLQHPHSHPRHPE 228
QY 84 --TSQALPAFTTEIOEASEGP 102
DB 229 KFPDKALXDNVCRNPDA SERP 249
RESULT 13
Q9PU78 PRELIMINARY; PRT; 313 AA.
ID Q9PU78;
AC Q9PU78;
DT 01-MAY-2000 (TREMBLrel. 13; Created)
DT 01-MAY-2000 (TREMBLrel. 13; Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21; Last annotation update)
DE Hepatocyte growth factor-like protein (Fragment).
OS Crocodylus niloticus (Nile crocodile) (African crocodile).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylia; Crocodyliidae; Crocodylus.
OX NCBI TaxID=8501;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER.
RX MEDLINE=20022983; PubMed=10555283;
RA Hughes S., Zetus D., Mouchiroud D.;
RT "Warm-blooded isochore structure in nile crocodile and turtle.";
RL Mol. Biol. Evol. 16:1521-1527(1999).
DR EMBL; AJ011396; CAB56422.1; -.
DR HSSP; P00747; 1HPJ.
DR MEROPS; S01.977; -.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Set_protease_Try.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
KW Hydrolyase; Serine protease.
FT NON TER 1
FT NON TER 1
FT SEQUENCE 313 AA; 34793 MW; 8B084704958B5AA2 CRC64;
SQ
Query Match 10.2%; Score 131.5; DB 13; Length 313;
Best Local Similarity 29.8%; Pred. No. 0.00079;
Matches 28; Conservative 11; Mismatches 42; Indels 13; Gaps 4

```

```
QY 4 CFWDNGHLVREDQTSAPGLRCLNWLDAQGLAS-APVSGAGNH--SYCRNPEDPRGP 59
DB 30 CYHNGELVYRGHTSKTRKGVTCQKWSQSPHVPQISTTHPAALIDENYCRNPNDSHGP 89
QY 60 WCYVSGE-----AGVPEKRPCEDLRCPELTSSQA 87
DB 90 WCYTDPRTPDYCGI---KPCAGDKIPSVLENA 120

RESULT 14
096SE8 PRELIMINARY; PRT; 154 AA.
ID 096SE8 01-DEC-2001 (Tremblrel. 19, Created)
AC 096SE8 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Urokinase-type plasminogen activator amino-terminal
fragment.
GN ATF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OK NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA Fu J., Bai X., Ruan C.;
RT "Cloning and expression of the amino-terminal fragment of human
RT urokinase-type plasminogen activator.";
RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RA Bai X., Fu J., Wang W., Xi X., Ruan C.;
RT "Overexpression of the amino-terminal fragment of human urokinase-type
RT plasminogen activator in breast cancer cells results in decreased
RT tumor invasion, growth and angiogenesis.";
RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY029537; AAK38734.1; -.
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00051; kringle_1.
DR InterPro; IPR000001; Kringle.
DR SMART; SM00181; EGF_1.
DR ProDom; PD000395; Kringle; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR Kase.
SQ SEQUENCE 154 AA; 17305 MW; A3CCP2FCFF50S572 CRC64;

Query Match 10.2%; Score 131; DB 4; Length 154;
Best Local Similarity 40.3%; Pred. No. 0.00036;
Matches 27; Conservative 7; Mismatches 25; Indels 8; Gaps 2;

QY 4 CFWDNGHLVREDQTSAPGLRCLNWLDAQGLASAPVSGAGNH--SYCRNPEDPRGP 56
DB 70 CTEGNGHFRKASITDTPMRPCLPWNASATVLTQTTTHAHRSDDLQGLGKMYICRNPD-NR 128
QY 57 RGPWCYV 63
DB 129 RRPWCYV 135

RESULT 15
09PTW7 PRELIMINARY; PRT; 608 AA.
ID 09PTW7 01-MAY-2000 (Tremblrel. 13, Created)
AC 09PTW7 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE Prothrombin.
GN OSPT.
OS Struthio camelus (Ostrich).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Palaeognathae; Struthioniformes; Struthionidae;
```

```
OC Struthio.
OK NCBI_TaxID=8801;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=20579470; PubMed=11137455;
RA Frost C., Nade R., Oelofsen W., Muramoto K., Naganuma T., Ogawa T.;
RT "Purification and characterization of ostrich prothrombin.";
RT Int. J. Biochem. Cell Biol. 32:1151-1159(2000).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AB028871; BAB9046.1; -.
DR HSSP; P00734; IUVS.
DR MEROPS; S01.217; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR002383; GLA blood.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser protease_Try.
DR InterPro; IPR000284; Vitr_dep_GLA.
DR Pfam; PF00594; gla; 1.
DR Pfam; PF00051; Kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00001; GLABLOOD.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00069; GLA; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Ser; 1.
DR PROSITE; PS00011; GH CARBOXYLATION; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS00240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
DR HydroLase; serine protease.
SQ SEQUENCE 608 AA; 69392 MW; 11B974B9ABE54EA2 CRC64;

Query Match 10.2%; Score 131; DB 13; Length 608;
Best Local Similarity 31.3%; Pred. No. 0.002;
Matches 41; Conservative 12; Mismatches 68; Indels 10; Gaps 5;

QY 2 GCGWDNGHLVREDQTSAPGLRCLNWLDAQGLASAPVSGAGN--HSYCRNPEDPRGP 59
DB 107 GNCSTDLGGVYRGHTSKTRKGVTCQKWSQSPHVPQISTTHPAALIDENYCRNPNDSHGP 166
QY 60 WCYVSGEAGVPEKRP---C-EDLRCPELTSSQALPAFTTEIOEASEG--PGADEVQVFAP 112
DB 167 WCYTRDPTVPRBESCFIPVCGERTTVEFTFRVAKPAPASTPCEGQKMLVAGTISVTI-SG 225
QY 113 ANLPAASENA 123
DB 226 AKCLPWNSEKA 236
```

Search completed: April 7, 2003, 09:21:35
Job time : 66.9023 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:24 ; Search time 45.2807 Seconds
(without alignments)
712.151 Million cell updates/sec

Title: US-10-057-951-2_COPY_22_263
Perfect score: 1289
Sequence: 1 SGGCFWMDNGHLXREDQTPA.....PVDPEGSTRPLMGAGTPGA 242

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1289	100.0	263	20	Kringlet protein s
2	1289	100.0	263	20	AAW05219
3	1289	100.0	263	22	AAW07769
4	1286	99.8	263	22	AAE00300
5	1280	99.3	263	21	AAU06149
6	1275	98.9	263	21	AAAB3237
7	1063.5	82.5	286	22	AAW03748
8	322	25.0	66	22	AAW05220
9	322	25.0	66	22	ABW37905
10	322	25.0	66	22	ABW37159
				22	AAW58537
					Human Brain expres

11	322	25.0	66	22	AAW71037	Human bone marrow
12	322	25.0	66	22	AAW18800	Peptide #5234 enco
13	322	25.0	66	22	AAW13114	Peptide #5351 enco
14	322	25.0	66	23	ABG40828	Human peptide enco
15	202	15.7	56	20	AAW12618	Human 5' EST seque
16	201	15.6	55	20	AAW12397	Human 5' EST seque
17	196	15.2	39	19	AAW72641	Nervous glia cell
18	192	14.9	39	19	AAW72640	Nervous glia cell
19	160.5	12.5	527	13	AAW20217	t-PA analogue expr
20	160.5	12.5	527	13	AAW20220	t-PA analogue expr
21	160.5	12.5	527	13	AAW20218	t-PA analogue expr
22	160.5	12.5	527	13	AAW20219	t-PA analogue expr
23	159.5	12.4	527	13	AAW20223	t-PA analogue expr
24	159.5	12.4	527	13	AAW20222	t-PA analogue expr
25	159.5	12.4	527	13	AAW20221	t-PA analogue expr
26	157.5	12.2	527	19	AAW54154	t-PA mutant (N142S
27	157.5	12.2	527	19	AAW54157	t-PA mutant (N142S
28	156.5	12.1	439	16	AAW68851	Delta 2-89 tissue
29	156.5	12.1	483	16	AAW70889	Human tissue PA va
30	156.5	12.1	483	16	AAW70887	Human tissue PA va
31	156.5	12.1	483	16	AAW70878	Human tissue PA va
32	156.5	12.1	483	16	AAW70879	Human tissue PA va
33	156.5	12.1	483	16	AAW70880	Human tissue PA va
34	156.5	12.1	483	16	AAW70881	Human tissue PA va
35	156.5	12.1	483	16	AAW70882	Human tissue PA va
36	156.5	12.1	483	16	AAW70883	Human tissue PA va
37	156.5	12.1	483	16	AAW70884	Human tissue PA va
38	156.5	12.1	483	16	AAW70885	Human tissue PA va
39	156.5	12.1	483	16	AAW70886	Human tissue PA va
40	156.5	12.1	483	16	AAW70887	Human tissue PA va
41	156.5	12.1	483	16	AAW70888	Human tissue PA va
42	156.5	12.1	483	16	AAW70890	Human tissue PA va
43	156.5	12.1	527	13	AAW21598	tPA variant - T103
44	156.5	12.1	527	14	AAW44812	Human tPA variant
45	156.5	12.1	527	14	AAW44816	Human tPA variant

ALIGNMENTS

RESULT 1	AAW05219	standard; Protein; 263 AA.
ID	AAW05219	standard; Protein; 263 AA.
AC	AAW05219	
DT	17-JUN-1999	(first entry)
XX		Kringlet protein sequence.
XX		Kringlet; cancer; inflammation; autoimmunity; rheumatoid arthritis; AIDS;
XX		CNS inflammation; cerebral degeneration; Alzheimer's disease; aschma;
XX		Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
XX		neurological abnormality; ischemia reperfusion injury; ischaemic injury;
KW		cardiovascular disease; kidney disease; liver disease; aplastic anaemia;
KW		myocardial infarction; hypotension; hypertension; allergy; infection;
KW		myelodysplastic syndrome; haematologic abnormality; diagnosis; therapy;
KW		male pattern baldness.
XX		
OS		Homo sapiens.
XX		
PN		WO9911788-A1.
XX		
PD		11-MAR-1999.
XX		
PF		02-SEP-1998; 98WO-US18270.
XX		
PR		01-SEP-1998; 98US-0144889.
XX		
PR		02-SEP-1997; 97US-0056032.
XX		
PA		(SMK) SMITHKLINE BEECHAM CORP.
XX		
PI		Albone EF, Kikly KK;

XX WPI: 1999-214707/18.
DR N-PSDB; MAX28354.
XX
XX New kringle1 polypeptides and polynucleotides
XX
XX Claim 1; Page 31-32; 42pp; English.
XX
CC This sequence is a kringle1 polypeptide of the invention.
CC The kringle1 polypeptides (I) are used to screen for agonists and
CC antagonists. Agonists are used to treat subjects in need of enhanced
CC activity or expression of (I). Antagonists are used to treat subjects
CC having need to inhibit the activity or expression of (I). The methods can
CC be used to treat conditions such as cancer, inflammation, autoimmunity,
CC allergy, asthma, rheumatoid arthritis, CNS inflammation, cerebellar
CC degeneration, Alzheimer's disease, Parkinson's disease, multiple
CC sclerosis, amyotrophic lateral sclerosis, head injury damage and other
CC neurological abnormalities, ischaemia reperfusion injury, cardiovascular
CC disease, kidney disease, liver disease, ischaemic injury, myocardial
CC infarction, hypertension, hypertension, AIDS, myelodysplastic syndromes
CC and other haematologic abnormalities, aplastic anaemia, male pattern
CC baldness, and bacterial, fungal, protozoan and viral infections. The
CC kringle1 polypeptides may also be used to generate antibodies.
CC Determining the presence or absence of mutations in, and analysing for
CC the presence or absence of expression of, kringle1 polynucleotides can be
CC used to diagnose a disease or susceptibility to a disease related to
CC expression or activity of kringle1 proteins. The polynucleotides may also
CC be used for chromosome identification, and mapping.
XX
SQ Sequence 263 AA;

Query Match 100.0%; Score 1289; DB 20; Length 263;
Best Local Similarity 100.0%; Pred. No. 4.6e-114;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGGCFWNGHLYREDTSPAPGIRCLNMLDAQSGLASAPVSGAGNHSYCRNPDEDPRGPW 60
DB 22 SGGCFWNGHLYREDTSPAPGIRCLNMLDAQSGLASAPVSGAGNHSYCRNPDEDPRGPW 81

QY 61 CVYSGAGVPEKPKPCEDLACPETTSQALPAFTTEIOEASEGAGDEVQFAPANALPARS 120
DB 82 CVYSGAGVPEKPKPCEDLACPETTSQALPAFTTEIOEASEGAGDEVQFAPANALPARS 141

QY 121 EAAAVQPVGISQRYVMSKEKKDLGTLGYVGIITMVTIINIGAGIILGYSYKKGXGLK 180
DB 142 EAAAVQPVGISQRYVMSKEKKDLGTLGYVGIITMVTIINIGAGIILGYSYKKGXGLK 201

QY 181 EGHDKVCEREMQRTITLPLSAFTNPICEIVDEKTVVHTSQTPVDPOEGSTPLMQAGTGP 240
DB 202 EGHDKVCEREMQRTITLPLSAFTNPICEIVDEKTVVHTSQTPVDPOEGSTPLMQAGTGP 261

QY 241 GA 242
DB 262 GA 263

RESULT 2
AAW87769
ID AAW87769 standard; Protein; 263 AA.
XX
AC AAW87769;
XX
XX 29-MAR-1999 (first entry)
XX
DE Human tissue plasminogen activator-like protease t-PALP.
XX
XX Tissue plasminogen activator-like protease; t-PALP; human;
KW circulatory system-related disorder; blood clotting; stroke;
KM thrombosis; peripheral arterial occlusion; pulmonary embolism;
KM myocardiothrombosis; diagnosis; therapy.
XX
XX Homo sapiens.
XX

FH Key Location/Qualifiers
FT Peptide 1..21
FT Protein /label= Sig_peptide
FT Protein 22..263
FT Domain /label= Mat_protein
FT Domain 25..84
FT Domain /note= "kringle domain"
FT Domain 85..263
FT Peptide /note= "protease domain"
FT Peptide 22..31
FT Peptide /note= "epitope-bearing region"
FT Peptide 35..44
FT Peptide /note= "epitope-bearing region"
FT Peptide 71..81
FT Peptide /note= "epitope-bearing region"
FT Peptide 91..107
FT Peptide /note= "epitope-bearing region"
FT Peptide 119..128
FT Peptide /note= "epitope-bearing region"
FT Peptide 138..147
FT Peptide /note= "epitope-bearing region"
FT Peptide 155..167
FT Peptide /note= "epitope-bearing region"
FT Peptide 193..203
FT Peptide /note= "epitope-bearing region"
FT Peptide 206..215
FT Peptide /note= "epitope-bearing region"
FT Peptide 227..237
FT Peptide /note= "epitope-bearing region"
FT Peptide 243..252
FT Peptide /note= "epitope-bearing region"

XX W09854199-A1.
XX
XX 03-DEC-1998.
XX
XX 27-MAY-1998; 98WO-US10728.
XX
XX 28-MAY-1997; 97US-0048000.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX
XX Ebnner R, Moore PA, Ruben SM;
XX
XX WPI: 1999-070207/06.
XX N-PSDB; AAV99636.
XX
XX New tissue plasminogen activator-like protease - useful in the
XX diagnosis and treatment of circulatory system-related disorders
XX
XX Claim 1; Page 56-57; 76pp; English.
XX
XX This is the amino acid sequence of tissue plasminogen activator-like
XX protease (t-PALP), a novel member of the serine protease family
XX that shares sequence homology to human tissue plasminogen activator
XX (see AAW87770). The t-PALP sequence was deduced from a cDNA clone
XX (see AAV99636) derived from activated monocytes. The 2.5 kb t-PALP
XX message has also been detected in heart, brain, lung, placenta,
XX liver, skeletal muscle, kidney, pancreas, spleen, thymus, prostate,
XX testis, ovary, small intestine, colon and peripheral blood
XX leukocytes. Isolated nucleic acids encoding amino acids -21 to
XX 242, -20 to 242, 1-242, 4-63 (kringle domain) and 64-242 (protease
XX domain) of t-PALP, or encoding epitope-bearing portions of t-PALP,
XX are also claimed, as are recombinant vectors, host cells, and
XX methods for producing t-PALP polypeptides. t-PALP may be used to
XX detect and treat disorders related to the circulatory system, and
XX to identify agonists and antagonists of t-PALP activity. The
XX homology between t-PALP and tPA indicates that t-PALP may be
XX involved in the regulation of normal and abnormal clotting
XX in e.g. stroke, deep-vein thrombosis, peripheral arterial
XX occlusion, pulmonary embolism and myocardiothrombosis.
XX
XX Sequence 263 AA;

	Query Match	Similarity	100.0%	Score 1289	DB 20	Length 263
	Best Local	Similarity	100.0%	Pred. No. 4.6e-114		
	Matches	242	Conservative	0	Mismatches	0
					Indels	0
					Gaps	0
Qy	1	SGCGFMDNCHLYREDQTSAPGRLCLNMTDAOSGLASAPVSGAGNHSYCRNPDEDPRGPW	60			
Db	22	SGCGFMDNCHLYREDQTSAPGRLCLNMTDAOSGLASAPVSGAGNHSYCRNPDEDPRGPW	81			
Qy	61	CYVSGAGVPEKRPCEDLDCPEFTTSQALPAFTTEIOEASEGGADEVQFAPANALPARS	120			
Db	82	CYVSGAGVPEKRPCEDLDCPEFTTSQALPAFTTEIOEASEGGADEVQFAPANALPARS	141			
Qy	121	EAAAVQPVIGISQRYVMSKEKKDGLTGIVLGIITMVTIIIAIGAGIILGYSYKRGKDXL	180			
Db	142	EAAAVQPVIGISQRYVMSKEKKDGLTGIVLGIITMVTIIIAIGAGIILGYSYKRGKDXL	201			
Qy	181	EGHDQVCEKREKQRIPLPSAFTNPCEIYDEKTVVYVHSQTPVDPQESTPLMGAGATP	240			
Db	202	EGHDQVCEKREKQRIPLPSAFTNPCEIYDEKTVVYVHSQTPVDPQESTPLMGAGATP	261			
Qy	241	GA 242				
Db	262	GA 263				

Accession	Protein	Location/Qualifiers
AAE00300	standard; Protein; 263 AA.	
AAE00300;		
13-JUN-2001	(first entry)	
Human tissue-plasminogen activator-like protease (t-PALP).		
Human; tissue-plasminogen activator-like protease; t-PALP; therapy; vascular disease; stroke; deep vein thrombosis; keloid; asthma; arterial occlusion; blood coagulation disorder; cerebroprotective; autoimmune system disorder; human immunodeficiency syndrome; cytostatic; rheumatoid arthritis; graft-versus-host disease; thyroiditis; cardiac; insulin dependent diabetes; inflammatory eye disease; allergy; antiviral; cardiovascular disease; heart disease; arrhythmia; myocardial ischaemia; hyperproliferative disorder; hypertrophic scar; neurological disease; Creutzfeldt-Jakob syndrome; neurodegenerative disorder; antiinflammatory; Alzheimer's disease; Parkinson's disease; immunosuppressive; antifungal; infectious disease; drug screening; gene therapy; neuroprotective; cancer; ophthalmological; antibacterial; vulnery.		
Homo sapiens.		
Key		
Binding-site	Location/Qualifiers	
	1..165	
	/note= "Binds to FLAG polypeptide to form t-PALP-FLAG fusion protein"	
Peptide	1..21	
	/label= Signal_peptide	
Domain	4..63	
	/label= Kringle_domain	
Region	12..21	
	/note= "Conserved region"	
Protein	22..263	
	/note= "Human mature tissue-plasminogen activator-like protease (t-PALP); Binds to FLAG polypeptide to form t-PALP-FLAG fusion protein"	
Region	22..38	
	/note= "Conserved region"	
Region	22..31	
	/note= "Epitope-bearing portion"	
Region	35..44	
	/note= "Epitope-bearing portion"	
Region	39..49	
	/note= "Conserved region"	

FT	Region	50..62	/note= "Conserved region"
FT	Region	63..84	/note= "Conserved region"
FT	Domain	64..242	/label= Protease_domain
FT	Region	71..81	/note= "Epitope-bearing portion"
FT	Region	85..97	/note= "Conserved region"
FT	Region	91..107	/note= "Epitope-bearing portion"
FT	Region	100..118	/note= "Conserved region"
FT	Region	119..128	/note= "Conserved region"
FT	Region	119..127	/note= "Conserved region"
FT	Region	128..143	/note= "Conserved region"
FT	Region	138..147	/note= "Conserved region"
FT	Region	146..163	/note= "Conserved region"
FT	Region	155..167	/note= "Conserved region"
FT	Region	164..180	/note= "Conserved region"
FT	Region	186..200	/note= "Conserved region"
FT	Region	193..203	/note= "Conserved region"
FT	Region	201..220	/note= "Conserved region"
FT	Region	206..215	/note= "Conserved region"
FT	Region	221..236	/note= "Conserved region"
FT	Region	227..237	/note= "Conserved region"
FT	Region	237..248	/note= "Conserved region"
FT	Region	243..252	/note= "Conserved region"
FT	Region	249..263	/note= "Conserved region"
FT	Region	/note= "Conserved region"	
XX	W0200125252-A1.		
XX	12-APR-2001.		
XX	PD		
XX	PP	03-OCT-2000; 2000MO-US27239.	
XX	PR	04-OCT-1999; 99US-0411977.	
XX	PA	(HUMA-) HUMAN GENOME SCI INC.	
PI	Moore PA, Ruben SM, Ehnert R;		
DR	WI; 2001-235402/24.		
DR	N-PSDB; AAD03460.		
XX			
XX			
PT	New (gene encoding and antibody immunospecific for a)		
PT	tissue-plasminogen activator-like protease, useful for the diagnosis		
PT	and treatment of (cardio)vascular diseases, hyperproliferative		
PT	disorders, immune system disorders and cancers -		
XX			
XX	Claim 17; Fig 1; 323pp; English.		
CC	The present amino acid sequence is HMSIB42 clone human		
CC	tissue-plasminogen activator-like protease (t-PALP). The t-PALP		
CC	sequence and their (ant)agonists are useful for the diagnosis and		
CC	treatment of vascular diseases e.g. stroke, deep vein thrombosis and		

Claim 17; Fig 1; 323pp; English.

arterial occlusion, blood coagulation disorders, (auto)immune system disorders e.g. human immunodeficiency syndrome, rheumatoid arthritis, graft-versus-host disease, thyroiditis, insulin dependent diabetes and inflammatory eye disease, allergic reactions e.g. asthma, cardiovascular diseases e.g. heart disease, arrhythmia and myocardial ischemia, hyperproliferative disorders, cancers, hypertrophic scars and keloids, neurological diseases e.g. Creutzfeldt-Jakob syndrome, neurodegenerative disorders e.g. Alzheimer's disease and Parkinson's disease and infectious disease e.g. viral, bacterial and fungal infections. The t-PALP sequences are also useful for drug screening. The t-PALP nucleotides are useful as chromosome markers and are involved in gene therapy.

Sequence 263 AA;

Query Match 100.0%; Score 1289; DB 22; Length 263;
Best Local Similarity 100.0%; Pred. No. 4,6e-114;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 SGGCFWNGHLYREDQTSAPGRLCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGPM 60
DB 22 SGGCFWNGHLYREDQTSAPGRLCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGPM 81
QY 61 CVYSGAGVPEKRPCEDLRCPEFTTSQALPAFTTEIQEASGSGADEVOYFAPANALPARS 120
DB 82 CVYSGAGVPEKRPCEDLRCPEFTTSQALPAFTTEIQEASGSGADEVOYFAPANALPARS 141
QY 121 EAAAVQPVIGISQVRNNSKEKKDLGTLGYVLGITMMVIIIAIGAGIILGYSKRGKDLK 180
DB 142 EAAAVQPVIGISQVRNNSKEKKDLGTLGYVLGITMMVIIIAIGAGIILGYSKRGKDLK 201
QY 181 EOHDOKVCEREMORITLPLSAFTNPCEIYDEKTVVHTSQTVPDQEGSTPLMGAGATP 240
DB 202 EOHDOKVCEREMORITLPLSAFTNPCEIYDEKTVVHTSQTVPDQEGSTPLMGAGATP 261
QY 241 GA 242
DB 262 GA 263

```

RESULT 4

AA086149
ID AA086149 standard; Protein; 263 AA.

```

XX AA086149;
AC 15-JUL-2002 (first entry)
DT 15-JUL-2002 (first entry)
XX Human PRO264 polypeptide.
DE Human; PRO; benign tumour; malignant tumour; lymphoid malignancy;
XX leukaemia; neuronal disorder; stromal disorder; blastocoealic disorder;
KW inflammatory disorder; immune disorder; angiogenic disorder;
KM cytostatic; neuroprotective.
XX Homo sapiens.
OS MO200153486-AA1.
PN 26-JUL-2001.
XX 26-JUL-2001.
PD 26-JUL-2001.
XX 11-FEB-2000; 2000WO-US03565.
PF 11-FEB-2000; 2000WO-US03565.
XX 08-MAR-1999; 99WO-US05028.
PR 11-MAR-1999; 99US-123972P.
XX 11-MAY-1999; 99US-133459P.
PR 02-JUN-1999; 99WO-US12252.
XX 22-JUN-1999; 99US-140650P.
PR 22-JUN-1999; 99US-140653P.
XX 26-JUL-1999; 99US-144758P.
PR 26-JUL-1999; 99US-145698P.
XX 28-JUL-1999; 99US-146222P.
PR 17-AUG-1999; 99US-149395P.
XX 31-AUG-1999; 99US-151689P.

```

```

PR 01-SEP-1999; 99WO-US02011.
PR 15-SEP-1999; 99WO-US21090.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 05-JAN-2000; 2000WO-US00219.
XX

```

(GENTH) GENENTECH INC.

PA Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
PI Masters SA, Pan J, Plichi RM, Roy MA, Smith V, Stone DM;
PI Matanabe CK, Wood WI;
XX

DR WPI; 2002-205567/26.

DR N-PSDB; ABR40275.

PT Thirty five nucleic acids encoding PRO polypeptides, useful for
PT treating benign or malignant tumours, leukaemias and lymphoid
PT malignancies, inflammatory, angiogenic and immunologic disorders -
XX Claim 61, Fig 44; 302pp; English.

XX The present invention relates to the isolation of novel human PRO
CC polypeptides and the polynucleotide sequences encoding them. The
CC PRO polypeptides, agonists, antagonists or anti-PRO antibodies are
CC useful for treating benign or malignant tumours (e.g. renal, kidney,
CC bladder, breast, etc), leukaemias and lymphoid malignancies, other
CC disorders such as neuronal, glial, astrocytal, hypothalamic, glandular,
CC macrophagal, stromal and blastocoealic disorders, inflammatory, immune
CC and angiogenic disorders. The polynucleotide sequences are also
CC useful in gene therapy. AA086128-AA086162 represent the human PRO
CC polypeptides of the invention.

Sequence 263 AA;

Query Match 99.8%; Score 1286; DB 23; Length 263;
Best Local Similarity 99.6%; Pred. No. 8,8e-114;
Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 SGGCFWNGHLYREDQTSAPGRLCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGPM 60
DB 22 SGGCFWNGHLYREDQTSAPGRLCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGPM 81
QY 61 CVYSGAGVPEKRPCEDLRCPEFTTSQALPAFTTEIQEASGSGADEVOYFAPANALPARS 120
DB 82 CVYSGAGVPEKRPCEDLRCPEFTTSQALPAFTTEIQEASGSGADEVOYFAPANALPARS 141
QY 121 EAAAVQPVIGISQVRNNSKEKKDLGTLGYVLGITMMVIIIAIGAGIILGYSKRGKDLK 180
DB 142 EAAAVQPVIGISQVRNNSKEKKDLGTLGYVLGITMMVIIIAIGAGIILGYSKRGKDLK 201
QY 181 EOHDOKVCEREMORITLPLSAFTNPCEIYDEKTVVHTSQTVPDQEGSTPLMGAGATP 240
DB 202 EOHDOKVCEREMORITLPLSAFTNPCEIYDEKTVVHTSQTVPDQEGSTPLMGAGATP 261
QY 241 GA 242
DB 262 GA 263

```

RESULT 5
AAB43237
ID AAB43237 standard; Protein; 263 AA.

```

XX AAB43237;
AC 08-FEB-2001 (first entry)
DT 08-FEB-2001 (first entry)
XX Human ORFX ORF3001 polypeptide sequence SEQ ID NO:6002.
DE Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnerable; antiparkinsonian; nocotropic; neuroprotective;
KM anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;

```

immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive; antiinflammatory; antiviral; antibacterial; antifungal; antineoplastic; antihypertensive; antianemic; gene therapy; cancer; proliferative disorder; hypertension; neurodegenerative disorder; osteoarthritis; graft vs host disease; cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS; cholesterol ester storage; systemic lupus erythematosus; infection; severe combined immunodeficiency; malaria; autoimmune disorder; aschma; allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound; bone damage; cartilage damage; antiinflammatory disease; coagulation; thrombosis; contraceptive.

OS Homo sapiens.

XX WO200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US08621.

XX 31-MAR-1999; 99US-0127607.

XX 02-APR-1999; 99US-0127636.

XX 05-APR-1999; 99US-0127728.

XX 30-MAR-2000; 2000US-0540763.

XX (CUBA-) CUBAGEN CORP.

XX Shinketsu RA, Leach M;

XX WPI; 2000-602362/57.

XX N-PSDB; AAC77446.

XX Novel nucleic acids and peptides derived from open reading frame X, useful for treating e.g. cancers, proliferative disorders, neurodegenerative disorders and cardiovascular disease -

XX Claim 11; Page 5181-5182; 5507pp; English.

CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397, which represent the human ORFX open reading frames 1 to 3161. The ORFX sequences have activities such as: cytostatic; hepatotropic; vulnary; antiproliferative; antiparkinsonian; nootropic; neuroprotective; osteoprotic; anticonvulsant; antirhectic; immunosuppressant; immunostimulant; cardiac; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive; antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic; antihypertoid; and antianemic. The sequences can be used for determining the presence of or predisposition to or preventing or treating pathological conditions associated with an ORFX-associated disorder. The nucleic acids can be used to express ORFX proteins in gene therapy vectors. The proteins and nucleic acids may be used to treat cancers, proliferative disorders, neurodegenerative disorders, osteoarthritis, graft vs host disease, cardiovascular disease, diabetes mellitus, hypertension, hypothyroidism, cholesterol ester storage, systemic lupus erythematosus, severe combined immunodeficiency (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to enhance coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 263 AA;

SO

Query Match 99.3%; Score 1280; DB 21; Length 263;

Best Local Similarity 99.2%; Pred. No. 3.3e-113;

Matches 240; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGGCFWMDNGHLVREDOTSPAPGRLCLNWLDAOSGLASAPVSGAGNHSCRNPDDEPRGPW 60

DB 22 SGGCFWMDNGHLVREDOTSPAPGRLCLNWLDAOSGLASAPVSGAGNHSCRNPDDEPRGPW 81

QY 61 CVVSGAGVPEKRPCEDLRCPELTTSQALPAFTTEIQEASESGADEVGVFAPANALPARS 120

DB 82 CVVSGAGVPEKRPCEDLRCPELTTSQALPAFTTEIQEASESGADEVGVFAPANALPARS 141

QY 121 EAAAVQPIGISORVRMNSKEKDLGTLGYLAGITMMVITIAAGIILGYSYRGKDLK 180

DB 142 EAAVQPIGISORVRMNSKEKDLGTLGYLAGITMMVITIAAGIILGYSYRGKDLK 201

QY 181 EQHDQKVCEREMQRTTLPISAFNPTCEIVDEKTVVHTSQTPVDPQSGSTPLMGQAGTP 240

DB 202 EQHDQKVCEREMQRTTLPISAFNPTCEIVDEKTVVHTSQTPVDPQSGSTPLMGQAGTP 261

QY 241 GA 242

DB 262 GA 263

RESULT 6

AAM93748

ID AAM93748 standard; Protein; 263 AA.

AC AAM93748;

XX 06-NOV-2001 (first entry)

DE Human polypeptide, SEQ ID NO: 3727.

XX Human, full length cDNA; cDNA synthesis; oligo-capping.

XX Homo sapiens.

XX EPI130094-A2.

XX 05-SEP-2001.

XX 07-JUL-2000; 2000EP-0114089.

XX 08-JUL-1999; 99EP-0194486.

XX 11-JAN-2000; 2000JP-0118774.

XX 02-MAY-2000; 2000JP-0183765.

XX (HELI-) HELIX RES INST.

XX Oca T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX WPI; 2001-524255/58.

DR N-PSDB; AAK94700.

XX 830 Primers useful for synthesizing full length cDNA clones and their use in genetic manipulation -

PT Claim 8; SEQ ID NO 3727; 1380pp + sequence listing; English.

PS The invention relates to primers for synthesizing full length cDNA clones. 830 cDNA molecules encoding a human protein have been isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have been determined. Primers for synthesizing the full length cDNA are useful for clarifying the function of the protein encoded by the cDNA. The full length clones were obtained by construction of full length enriched cDNA libraries that were synthesised by the oligo-capping method. The primers enable the production of the full length cDNA easily without any special methods. The present sequence is a polypeptide encoded by a full length human cDNA of the invention.

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in CD-ROM format directly from EPO.

XX Sequence 263 AA;

SO

Query Match 98.9%; Score 1275; DB 22; Length 263;

Best Local Similarity 99.2%; Pred. No. 9.8e-113;

Matches 240; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SGGCFWMDNGHLVREDOTSPAPGRLCLNWLDAOSGLASAPVSGAGNHSCRNPDDEPRGPW 60

DB 22 SGGCFWMDNGHLVREDOTSPAPGRLCLNWLDAOSGLASAPVSGAGNHSCRNPDDEPRGPW 81

QY 61 CYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASBEPGADDEVQVFPANALPARS 120
 DB 82 CYSGEAGVPEKRPCEGRCPETTSQALPAFTTEIQEASBEPGADDEVQVFPANALPARS 141
 QY 121 EAAAVQPIVIGISQVRNMSKEKKDGLGVLGVTMWVITIIAIGAGIILGYSYRGKDLK 180
 DB 142 EAAAVQPIVIGISQVRNMSKEKKDGLGVLGVTMWVITIIAIGAGIILGYSYRGKDLK 201
 QY 181 EQHDKVCEREMQRTITLPLSAFTNPTCEIVDEKTVVHTSQTVPDPOEGSTPLMGQAGTP 240
 DB 202 EQHDKVCEREMQRTITLPLSAFTNPTCEIVDEKTVVHTSQTVPDPOEGSTPLMGQAGTP 261
 QY 241 GA 242
 DB 262 GA 263

RESULT 7
 ID AAY05220 standard; Protein; 286 AA.
 AC AAY05220;
 DT 17-JUN-1999 (first entry)
 DE Kringle1 protein sequence.
 KM Kringle1; cancer; inflammation; autoimmunity; rheumatoid arthritis; AIDS;
 KM CNS inflammation; cerebellar degeneration; Alzheimer's disease; asthma;
 KM Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
 KM neurological abnormality; ischemia reperfusion injury; ischaemic injury;
 KM cardiovascular disease; kidney disease; liver disease; aplastic anaemia;
 KM myocardial infarction; hypotension; hypertension; allergy; infection;
 KM myelodysplastic syndrome; haematologic abnormality; diagnosis; therapy;
 KM male pattern baldness.
 OS Homo sapiens.
 PN WO9911788-A1.
 PD 11-MAR-1999.
 PF 02-SEP-1998; 98WO-US18270.
 PR 01-SEP-1998; 98US-0144889.
 PR 02-SEP-1997; 97US-0056032.
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PI Albione EF, Kikly KK;
 DR WPI; 1999-214707/18.
 DR N-PSDB; AAX28355.
 PT New Kringle1 polypeptides and polynucleotides
 PS Claim 14; Page 33; 42pp; English.

This sequence is a Kringle1 polypeptide of the invention.
 The Kringle1 polypeptides (I) are used to screen for agonists and antagonists. Agonists are used to treat subjects in need of enhanced activity or expression of (I). Antagonists are used to treat subjects having need to inhibit the activity or expression of (I). The methods can be used to treat conditions such as cancer, inflammation, autoimmunity, allergy, asthma, rheumatoid arthritis, CNS inflammation, cerebellar degeneration, Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis, head injury, damage and other neurological abnormalities, ischemia reperfusion injury, cardiovascular disease, kidney disease, liver disease, ischaemic injury, myocardial infarction, hypotension, hypertension, AIDS, myelodysplastic syndromes and other haematologic abnormalities, aplastic anaemia, male pattern baldness, and bacterial, fungal, protozoan and viral infections. The

CC Kringle1 polypeptides may also be used to generate antibodies.
 CC Determining the presence or absence of mutations in, and analyzing for
 CC the presence or absence of expression of, Kringle1 polynucleotides can be
 CC used to diagnose a disease or susceptibility to a disease related to
 CC expression or activity of Kringle1 proteins. The polynucleotides may also
 CC be used for chromosome identification, and mapping.

SO Sequence 286 AA:
 Query Match 82.5%; Score 1063.5; DB 20; Length 286;
 Best Local Similarity 86.2%; Pred. No. 1.3e-92;
 Matches 206; Conservative 5; Mismatches 27; Indels 1; Gaps 1;

QY 1 SGGCFMNDGNHLYREDQTSAPAGRCINWLDASGASAPVSGAGNHSYCRNDEDPGRPW 60
 DB 22 SGGCFMNDGNHLYREDQTSAPAGRCINWLDASGASAPVSGADHSYCRNDEDPGRPW 81
 QY 61 CYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASBEPGADDEVQVFPANALPAR 119
 DB 82 CYSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASBEPGADDEVQVFPANALPAR 141
 QY 120 SEAAVQPIVIGISQVRNMSKEKKDGLGVLGVTMWVITIIAIGAGIILGYSYRGKDL 179
 DB 142 SEAAVQPIVIGISQVRNMSKEKKDGLGVLGVTMWVITIIAIGAGIILGYSYRGKDL 201
 QY 180 KEQHDQVCEREMQRTITLPLSAFTNPTCEIVDEKTVVHTSQTVPDPOEGSTPLMGQAG 238
 DB 202 KEQHDQVCEREMQRTITLPLSAFTNPTCEIVDEKTVVHTSQTVPDPOEGSTPLMGQAG 260

RESULT 8
 ID ABB37905 standard; Peptide; 66 AA.
 AC ABB37905;
 DT 04-FEB-2002 (first entry)
 DE Peptide #5411 encoded by human foetal liver single exon probe.
 KM Human; foetal liver; gene expression; single exon nucleic acid probe.
 OS Homo sapiens.
 PN WO200157277-A2.
 PD 09-AUG-2001.
 PF 30-JAN-2001; 2001WO-US00669.
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI; 2001-483447/52.
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human foetal liver -
 PS Claim 27; SEQ ID NO 30540; 639pp + sequence listing; English.

The invention relates to a single exon nucleic acid probe for measuring human gene expression in a sample derived from human foetal liver. The single exon nucleic acid probes may be used for predicting, measuring and displaying gene expression in samples derived from human

CC fetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 66 AA;

Query Match 25.0%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 4,4e-23;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 TTSQLPFTTETIQASGPGADEVOVFAPANALPARSEAAAVQPVIGISQRYMNSKEK 142
Db 1 TTSQLPFTTETIQASGPGADEVOVFAPANALPARSEAAAVQPVIGISQRYMNSKEK 60
Qy 143 KDLGTL 148
Db 61 KDLGTL 66

RESULT 9
ABR23159
ID ABR23159 standard; Protein; 66 AA.
XX
AC ABR23159;
XX
DT 23-JAN-2002 (first entry)
XX
DE Protein #5158 encoded by probe for measuring heart cell gene expression.
XX
KW Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
OS Homo sapiens.
XX
PN MO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001MO-US00666.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
XX
PT hearts -
XX
PS Claim 15; SEQ ID No 24929; 530pp; English.
XX
XX

CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA2153-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting, the
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 66 AA;

Query Match 25.0%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 4,4e-23;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 TTSQLPFTTETIQASGPGADEVOVFAPANALPARSEAAAVQPVIGISQRYMNSKEK 142
Db 1 TTSQLPFTTETIQASGPGADEVOVFAPANALPARSEAAAVQPVIGISQRYMNSKEK 60
Qy 143 KDLGTL 148
Db 61 KDLGTL 66

RESULT 10
AAM58537
ID AAM58537 standard; Protein; 66 AA.
XX
AC AAM58537;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 30642.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN MO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001MO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
XX
PT brains -
XX
PS Example 4; SEQ ID NO: 30642; 650pp + Sequence Listing; English.
XX
XX

CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
XX

Qy 83 TTSQLPFTTETIQASGPGADEVOVFAPANALPARSEAAAVQPVIGISQRYMNSKEK 142

Query Match 25.0%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 4,4e-23;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Db      1  TTSQALPAFTTEIQASBGPGADEVOVFAPANALPARSEAAAVQPIGISOVRNMSKEX 60
QY      143 KDLGTL 148
XX      |||||
DT      61 KDLGTL 66

Db      61 KDLGTL 66

RESULT 11
AAM71037
ID      AAM71037 standard; Protein; 66 AA.
XX
AC      AAM71037;
XX
DT      06-NOV-2001 (first entry)
XX
DE      Human bone marrow expressed probe encoded protein SEQ ID NO: 31343.
XX
KW      Human; bone marrow expressed exon; gene expression analysis; probe;
KW      microarray; cancer; leukemia; lymphoma; myeloma.
XX
OS      Homo sapiens.
XX
PN      WO200157276-A2.
XX
PD      09-AUG-2001.
XX
PF      30-JAN-2001; 2001WO-US00668.
XX
PR      04-FEB-2000; 2000US-0180312.
PR      26-MAY-2000; 2000US-0207456.
PR      30-JUN-2000; 2000US-0608408.
PR      03-AUG-2000; 2000US-0632366.
PR      21-SEP-2000; 2000US-0234687.
PR      27-SEP-2000; 2000US-0236359.
PR      04-OCT-2000; 2000GB-0024263.
XX
PA      (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI      Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR      WPI; 2001-488900/53.
XX
PT      Human genome-derived single exon nucleic acid probes useful for
PT      analyzing gene expression in human bone marrow -
XX
PS      Example 4; SEQ ID NO: 31343; 658bp + Sequence listing; English.
XX
CC      The present invention provides a number of single exon nucleic acid
CC      probes which are derived from genomic sequences expressed in the human
CC      bone marrow. They can be used to measure gene expression in bone marrow
CC      samples, which may enable the improved diagnosis and treatment of cancers
CC      such as lymphoma, leukemia and myeloma. The present sequence is a
CC      protein encoded by one of the probes of the invention.
XX
SQ      Sequence 66 AA;

Query Match      25.0%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 4.4e-23;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      83 TTSQALPAFTTEIQASBGPGADEVOVFAPANALPARSEAAAVQPIGISOVRNMSKEX 142
DB      1 TTSQALPAFTTEIQASBGPGADEVOVFAPANALPARSEAAAVQPIGISOVRNMSKEX 60
QY      143 KDLGTL 148
DB      61 KDLGTL 66

RESULT 12,
AAM18800
ID      AAM18800 standard; Protein; 66 AA.

```

```

XX      AAM18800;
AC      12-OCT-2001 (first entry)
XX
DT      Peptide #5234 encoded by probe for measuring cervical gene expression.
XX
DE      Probe; human; microarray; gene expression; cervical epithelial cell;
KW      cervical cancer.
XX
OS      Homo sapiens.
XX
PN      WO200157278-A2.
XX
PD      09-AUG-2001.
XX
PF      30-JAN-2001; 2001WO-US00670.
XX
PR      04-FEB-2000; 2000US-0180312.
PR      26-MAY-2000; 2000US-0207456.
PR      30-JUN-2000; 2000US-0608408.
PR      03-AUG-2000; 2000US-0632366.
PR      21-SEP-2000; 2000US-0234687.
PR      27-SEP-2000; 2000US-0236359.
PR      04-OCT-2000; 2000GB-0024263.
XX
PA      (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI      Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR      WPI; 2001-488901/53.
XX
PT      Human genome-derived single exon nucleic acid probes useful for
PT      analyzing gene expression in human cervical epithelial cells -
XX
PS      Claim 27; SEQ ID No 23626; 487bp; English.
XX
CC      The present invention relates to human single exon nucleic acid probes
CC      (SENP; see A110068-A126459). The present sequence is a peptide encoded
CC      by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC      can be used to produce a single exon microarray, which can be used for
CC      measuring human gene expression in a sample derived from human cervical
CC      epithelial cells. By measuring gene expression, the probes are therefore
CC      useful in grading and/or staging of diseases of the cervix, notably
CC      cervical cancer.
CC      Note: The sequence data for this patent did not form part of the printed
CC      specification, but was obtained in electronic format directly from WIPO
CC      at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ      Sequence 66 AA;

Query Match      25.0%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 4.4e-23;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      83 TTSQALPAFTTEIQASBGPGADEVOVFAPANALPARSEAAAVQPIGISOVRNMSKEX 142
DB      1 TTSQALPAFTTEIQASBGPGADEVOVFAPANALPARSEAAAVQPIGISOVRNMSKEX 60
QY      143 KDLGTL 148
DB      61 KDLGTL 66

RESULT 13
AAM31314
ID      AAM31314 standard; Protein; 66 AA.
XX
AC      AAM31314;
XX
DT      17-OCT-2001 (first entry)
XX
DE      Peptide #5351 encoded by probe for measuring placental gene expression.

```

XX Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200157272-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00663.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI, 2001-488897/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -
 XX
 PS Claim 27; SEQ ID No 31583; 654bp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes (SENP;
 CC see A431315-A4157546). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders.
 CC
 XX
 SQ Sequence 66 AA;
 Query Match 25.0%; Score 322; DB 22; Length 66;
 Best Local Similarity 100.0%; Pred. No. 4.4e-23;
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 83 TTGQALPFTTETIQASGPGADVEVPAPANNALPASEAAAVPYTIGISQVRMNSKEX 142
 DB 1 TTGQALPFTTETIQASGPGADVEVPAPANNALPASEAAAVPYTIGISQVRMNSKEX 60
 QY 143 KDLGTL 148
 DB 61 KDLGTL 66
 RESULT 14
 ABG40828
 ID ABG40828 standard; Peptide; 66 AA.
 XX
 AC ABG40828;
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE Human peptide encoded by genome-derived single exon probe SEQ ID 30493.
 XX
 KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemorrhoidosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX

OS Homo sapiens.
 XX
 PN WO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00665.
 XX
 PR 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI, 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 XX
 PS Claim 27; SEQ ID No 30493; 634bp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes. The novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung, measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectable
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemorrhoidosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karsenger syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 66 AA;
 Query Match 25.0%; Score 322; DB 23; Length 66;
 Best Local Similarity 100.0%; Pred. No. 4.4e-23;
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 TTSGALPATTTHIQASRGPGADEVQVFPAPNALPARSEAAVQVITGISQVRNMSKEK 142
 DB 1 TTSGALPATTTHIQASRGPGADEVQVFPAPNALPARSEAAVQVITGISQVRNMSKEK 60
 QY 143 KDLGTL 148
 DB 61 KDLGTL 66

RESULT 15

AAV12615
 ID AAV12615 standard; Protein; 56 AA.

XX AAV12615;

DT 22-JUN-1999 (first entry)

DE Human 5' EST secreted protein SEQ ID NO: 280 from WO 9906553.

XX Human; secreted protein; EST; expressed sequence tag; diagnosis;

KM forensic; gene therapy; chromosome mapping; signal peptide;

KM upstream regulatory sequence; cytokine activity; cell proliferation;

KM differentiation; haematopoiesis regulation; tissue growth regulation;

KM reproductive hormone regulation; chemoketic; haemostatic;

KM thrombolytic; antiinflammatory; tumour inhibition; antitumour.

XX Homo sapiens.

XX WO9906553-A2.

XX 11-FEB-1999.

PF 31-JUL-1998; 98WO-IB01237.

XX 01-AUG-1997; 97US-0905051.

XX (GSEST) GENSET.

PI DucJert A, Dumas Milne Edwards J, Lacroix B;

DR WPI; 1999-153783/13.

XX N-PSDB; AAX41473.

XX New nucleic acids encoding human secreted proteins - obtained from

PT CDNA libraries derived from umbilical cord, lymph ganglia,

PT lymphocytes and placental tissue

XX Claim 34; Page 376; 411pp; English.

XX AAX41379 to AAX41526 represent 5' expressed sequence tags (ESTs) for

CC human secreted proteins, and encode the proteins given in AAV12521 to

CC AAV12668, respectively. The proteins given represent the signal peptide

CC and an N-terminal fragment of a secreted protein. The nucleic acid

CC sequences can be used for producing secreted human gene products. They

CC can also be used to develop products for diagnosis and therapy. The

CC protein obtained may have cytokine activity, cell

CC proliferation/differentiation activity, haematopoiesis regulating

CC activity, tissue growth regulation activity, reproductive hormone

CC regulating activity, chemotactic/chemokinetic activity, haemostatic and

CC thrombolytic activity, receptor/ligand activity, antiinflammatory

QY 1 SGGCFWNGHLYREDQTSAPGLRCLNWLDAOSGL 35
 DB 22 SGGCFWNGHLYREDQTSAPGLRCLNWLDAOSGL 56
 Search completed: April 7, 2003, 09:18:32
 Job time : 45.2807 secs

Query Match 15.7%; Score 202, DB 20; Length 56;
 Best Local Similarity 97.1%; Pred. No. 8.7e-12;
 Matches 34; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:19:28 ; Search time 17.106 Seconds

(without alignments)
864,895 Million cell updates/sec

Title: US-10-057-951-2_COPY_22_263

Perfect score: 1289

Sequence: 1 SGGCFWMDNGHLYREDQTSFA.....PVDPQEGSTPLMGAGTPGA 242

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1289	100.0	263	9	US-10-102-704-2
2	1289	100.0	263	9	US-10-057-951-2
3	1289	100.0	263	10	US-09-084-491A-2
4	1289	100.0	263	10	US-09-084-491A-2
5	154.5	12.0	527	9	US-09-987-457-18
6	154.5	12.0	527	9	US-09-987-457-18
7	154.5	12.0	527	9	US-09-987-457-18
8	154.5	12.0	527	9	US-09-987-457-18
9	154.5	12.0	527	9	US-09-987-457-18
10	142	11.0	160	9	US-09-335-325-35
11	142	11.0	160	9	US-09-335-325-35
12	142	11.0	160	9	US-09-335-325-35
13	142	11.0	160	9	US-09-335-325-35
14	142	11.0	160	9	US-09-335-325-35
15	142	11.0	250	9	US-10-131-241-30
16	142	11.0	250	9	US-10-131-241-30
17	142	11.0	250	9	US-10-131-241-30
18	142	11.0	250	9	US-10-131-241-30
19	142	11.0	250	9	US-10-131-241-30

20	142	11.0	352	9	US-09-335-325-40	Sequence 40, Appl
21	142	11.0	352	9	US-10-131-241-40	Sequence 40, Appl
22	142	11.0	352	10	US-09-761-120-40	Sequence 40, Appl
23	142	11.0	358	10	US-09-761-120-40	Sequence 42, Appl
24	142	11.0	378	9	US-09-335-325-42	Sequence 42, Appl
25	142	11.0	378	9	US-10-131-241-42	Sequence 42, Appl
26	142	11.0	458	10	US-09-873-676-1	Sequence 4, Appl
27	142	11.0	458	10	US-09-873-676-1	Sequence 4, Appl
28	142	11.0	559	10	US-09-946-893-4	Sequence 5, Appl
29	142	11.0	559	10	US-09-946-893-4	Sequence 5, Appl
30	142	11.0	576	10	US-09-946-893-6	Sequence 6, Appl
31	142	11.0	791	9	US-09-967-386-1	Sequence 1, Appl
32	142	11.0	810	10	US-09-967-386-1	Sequence 1, Appl
33	138	10.7	79	9	US-09-335-325-8	Sequence 8, Appl
34	138	10.7	79	9	US-10-131-241-8	Sequence 8, Appl
35	138	10.7	79	10	US-09-753-064-2	Sequence 2, Appl
36	138	10.7	79	10	US-09-753-064-2	Sequence 2, Appl
37	132	10.2	411	10	US-09-880-503-3	Sequence 3, Appl
38	132	10.2	411	10	US-09-880-503-3	Sequence 3, Appl
39	132	10.2	431	9	US-10-076-421-2	Sequence 1, Appl
40	131.5	10.2	96	10	US-09-264-468B-1	Sequence 9, Appl
41	131.5	10.2	143	10	US-09-880-503-8	Sequence 8, Appl
42	131.5	10.2	88	10	US-09-880-503-8	Sequence 1, Appl
43	131	10.2	135	10	US-09-880-503-4	Sequence 1, Appl
44	131	10.2	138	10	US-10-237-667-12	Sequence 12, Appl
45	131	10.2	138	9	US-10-237-708-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-10-102-704-2
Sequence 2, Application US/10102704
Patent No. US020020164768A1
GENERAL INFORMATION:
APPLICANT: Moore et al.
TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protein
FILE REFERENCE: P9378C1
CURRENT APPLICATION NUMBER: US/10/102,704
CURRENT FILING DATE: 2002-03-22
PRIOR APPLICATION NUMBER: 09/084,491
PRIOR FILING DATE: 1998-05-27
PRIOR APPLICATION NUMBER: 60/048,000
PRIOR FILING DATE: 1997-05-28
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
US-10-102-704-2

Query Match 100.0%; Score 1289; DB 9; Length 263;
Best Local Similarity 100.0%; Pred. No. 4e-110;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGGCFWMDNGHLYREDQTSFA...PVDPQEGSTPLMGAGTPGA 60
Db 22 SGGCFWMDNGHLYREDQTSFA...PVDPQEGSTPLMGAGTPGA 81
QY 61 CYVSGAGVPEKPCEDRCPTTSOLPAFTTEIOEASBEGGADDEVQVAPNALPARS 120
Db 82 CYVSGAGVPEKPCEDRCPTTSOLPAFTTEIOEASBEGGADDEVQVAPNALPARS 141
QY 121 EAAAVOPVIGISORVRNMSKEXKDI...GTVGIVGIMVYIIAIGAGIILGYSYKRXKDK 180
Db 142 EAAAVOPVIGISORVRNMSKEXKDI...GTVGIVGIMVYIIAIGAGIILGYSYKRXKDK 201
QY 181 EHDQVCEREMQRTTLP...LAFNPFCEIYDEKTVVHTSQT...PVDPQEGSTPLMGAGTPGA 240
Db 202 EHDQVCEREMQRTTLP...LAFNPFCEIYDEKTVVHTSQT...PVDPQEGSTPLMGAGTPGA 261

QY 241 GA 242
DB 262 GA 263

RESULT 2

US-10-057-951-2
; Sequence 2, Application US/10057951
; Patent No. US2002017213A1
; GENERAL INFORMATION:
; APPLICANT: Moore et al.
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
; FILE REFERENCE: P378P1
; CURRENT APPLICATION NUMBER: US/10/057,951
; PRIORITY FILING DATE: 1997-05-29
; PRIORITY FILING DATE: 1999-10-04
; PRIORITY FILING DATE: 1998-05-27
; PRIORITY FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-951-2

Query Match

Best Local Similarity 100.0%; Score 1289; DB 9; Length 263;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGGCFMNDGHLVREDQTSAPGLRCINWLDASGLASAPVSGAGNHSYCRNPDEDPGFW 60
DB 22 SGGCFMNDGHLVREDQTSAPGLRCINWLDASGLASAPVSGAGNHSYCRNPDEDPGFW 81
QY 61 CYSVSGAVPEKRPCEIDRCPTTSQALPAFTTEIOEASBEGGADEVQVFPANALPARS 120
DB 82 CYSVSGAVPEKRPCEIDRCPTTSQALPAFTTEIOEASBEGGADEVQVFPANALPARS 141
QY 121 EAAAVOPVIGISQVRNMSKEKDLGTLGVGLGTTMVTIIAIGAGIILGYSYRGKDLK 180
DB 142 EAAAVOPVIGISQVRNMSKEKDLGTLGVGLGTTMVTIIAIGAGIILGYSYRGKDLK 201
QY 181 EOHDOVCEREMQRTLPISAFNPTCEIYDEKTVVHTSQTPVDPQSGSTPLMGQAGTP 240
DB 202 EOHDOVCEREMQRTLPISAFNPTCEIYDEKTVVHTSQTPVDPQSGSTPLMGQAGTP 261
QY 241 GA 242
DB 262 GA 263

RESULT 3

US-09-084-491A-2
; Sequence 2, Application US/09084491A
; Patent No. US20020061576A1
; GENERAL INFORMATION:
; APPLICANT: MOORE, PAUL A.
; APPLICANT: RUBEN, STEVEN M.
; APPLICANT: BENNER, REINHARD
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/084,491A
FILING DATE: 27-MAY-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: P378
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-084-491A-2

Query Match 100.0%; Score 1289; DB 10; Length 263;
Best Local Similarity 100.0%; Pred. No. 4e-110;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGGCFMNDGHLVREDQTSAPGLRCINWLDASGLASAPVSGAGNHSYCRNPDEDPGFW 60
DB 22 SGGCFMNDGHLVREDQTSAPGLRCINWLDASGLASAPVSGAGNHSYCRNPDEDPGFW 81
QY 61 CYSVSGAVPEKRPCEIDRCPTTSQALPAFTTEIOEASBEGGADEVQVFPANALPARS 120
DB 82 CYSVSGAVPEKRPCEIDRCPTTSQALPAFTTEIOEASBEGGADEVQVFPANALPARS 141
QY 121 EAAAVOPVIGISQVRNMSKEKDLGTLGVGLGTTMVTIIAIGAGIILGYSYRGKDLK 180
DB 142 EAAAVOPVIGISQVRNMSKEKDLGTLGVGLGTTMVTIIAIGAGIILGYSYRGKDLK 201
QY 181 EOHDOVCEREMQRTLPISAFNPTCEIYDEKTVVHTSQTPVDPQSGSTPLMGQAGTP 240
DB 202 EOHDOVCEREMQRTLPISAFNPTCEIYDEKTVVHTSQTPVDPQSGSTPLMGQAGTP 261
QY 241 GA 242
DB 262 GA 263

RESULT 4

US-09-864-761-38457
; Sequence 38457, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wenheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIORITY FILING DATE: 2001-05-23
; PRIORITY FILING DATE: 2000-02-04
; PRIORITY FILING DATE: 2000-05-26
; PRIORITY FILING DATE: US 09/632,366
; PRIORITY FILING DATE: 2000-08-03
; PRIORITY FILING DATE: GB 24263.6
; PRIORITY FILING DATE: 2000-10-04
; PRIORITY FILING DATE: US 60/236,359
; PRIORITY FILING DATE: 2000-09-27
; PRIORITY FILING DATE: PCT/US01/00666
; PRIORITY FILING DATE: 2001-01-30

```

Query Match: 25.0%; Score 322; DB 10; Length 66;
Best Local Similarity 100.0%; Pred. No. 1.4e-22;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      83  TTSQALPAFTTEIQEASGEGCADEVQVFAFPANMLPARSEAAAVQPIIGISQRYRNNSEK 142
Db       1  TTSQALPAFTTEIQEASGEGCADEVQVFAFPANMLPARSEAAAVQPIIGISQRYRNNSEK 60
          |||||
QY      143  KDIGTL 148
Db       61  KDIGTL 66

RESULT 5
; Sequence 18, Application US/09987457
; Publication No. US20030013150A1
; GENERAL INFORMATION:
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tayapiwatana, Chatchai
; APPLICANT: Goetzi, Friedrich
; APPLICANT: Werner, Rolf-Guenther
; TITLE OR INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 065.2180001
; CURRENT APPLICATION NUMBER: US/09/987,457
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,573
; PRIOR FILING DATE: 2001-02-15

```

Query Match	12.0%	Score 154.5	DB 9	Length 527
Best Local Similarity	39.6%	Pred. 10.4e-06		
Matches	36	Conservative	6	Mismatches 36; Indels 13; Gaps 4
QY	4	CFENDGNHLYREDCSPAPRGLCNMLDNOGLAAPS-----		-GAGNYSYCRNDE 54
DB	92	CIDDOGISYRTGWTATASGAELTNW--NSSALAQKPSYSGRRPADRLTGLGNHNYCRPDR		149
QY	55	DPKGPWCYSYSGAGVPEKRPCECDLRCPETTS		85
DB	150	DSK-PWCYVF-KAGKYSSEFCSTPAASGNS		178

```

Query Match 12.0%; Score 154.5; DB 9; Length 527;
Best Local Similarity 39.6%; Pred. No. 4,4e-06;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4

Oy 4 CPMDNHLVREDOEPPAPGLRCLNMLDAOSGLASAPVS-----GAGNHSYCRANDE 54
Db 92 CYBDQGISRYGTWSTVASSGAGACTNW--NSSLAQKPSGRRPDAIRLGIQNNYCRNDR 149

Oy 55 DPRGWCYVSGEAGVPEKRPCEDLRCEPTTS 85
Db 150 DSK-FWCYVF-KAGVYSFECSTPASEGNS 178

RESULT 7
US-09-974-298-145
; Sequence 145, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hui-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04

```

PRIOR APPLICATION NUMBER: 60/238,331
PRIOR FILING DATE: 2000-05-10
NUMBER OF SEQ ID NOS: 194
SOFTWARE: PERL Program
SEQ ID NO 145
LENGTH: 562
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20020156263A1 1001470CD1
US-09-974-298-145

Query Match 12.0%; Score 154.5; DB 9; Length 562;
Best Local Similarity 39.6%; Pred. No. 4,7e-06;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 4 CFWDNGHLVREDQTSAPGLRCLNWLDAQSGIASAPVS-----GAGNHSYCRNPDE 54
DB 127 CYEDOGISYRGTSWSTAESGAECTNW--NSSALAQKPYSGRRPDALRLGNGHNYCRNPDR 184

QY 55 DPGPWCYVSGEAGVPEKRPCEDLRCPEPTS 85
DB 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 213

RESULT 8
US-09-969-271-7
Sequence 7, Application US/09969271
Patent No. US20020098179A1
GENERAL INFORMATION:
APPLICANT: Pfizer Inc. (A1) designated States except GB and EP (GB);
APPLICANT: Pfizer Limited (GB and EP (GB) only)
TITLE OF INVENTION: Pharmaceutical Combinations
FILE REFERENCE: PCS10951APME
CURRENT APPLICATION NUMBER: US/09/969,271
CURRENT FILING DATE: 2001-10-01
PRIOR APPLICATION NUMBER: GB 0025473.0
PRIOR FILING DATE: 2000-10-17
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PasteSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 562
TYPE: PRT
ORGANISM: Homo sapiens
US-09-969-271-7

Query Match 12.0%; Score 154.5; DB 10; Length 562;
Best Local Similarity 39.6%; Pred. No. 4,7e-06;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 4 CFWDNGHLVREDQTSAPGLRCLNWLDAQSGIASAPVS-----GAGNHSYCRNPDE 54
DB 127 CYEDOGISYRGTSWSTAESGAECTNW--NSSALAQKPYSGRRPDALRLGNGHNYCRNPDR 184

QY 55 DPGPWCYVSGEAGVPEKRPCEDLRCPEPTS 85
DB 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 213

RESULT 9
US-09-335-325-35
Sequence 35, Application US/09335325
Patent No. US20020164717A1
GENERAL INFORMATION:
APPLICANT: Folkman, M. Judah
O'Reilly, Michael
Cao, Yihai
Sim, B. Kim Lee
TITLE OF INVENTION: Angiostatin Fragments and Method of Use
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSER: Jones & Askeew

STREET: 191 Peachtree Street, 37th Floor
CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 09/335,325
FILING DATE: 17-Jun-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/612,788
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Warren, William L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 05213-0126
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 160 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: K1-2
SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-335-325-35

Query Match 11.0%; Score 142; DB 9; Length 160;
Best Local Similarity 37.2%; Pred. No. 1.3e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

QY 4 CFWDNGHLVREDQTSAPGLRCLNWLDA-----QSGIASAPVSGAGNHSYCRNPDEDPG 58
DB 1 CKTGKGYRGTSWSTAESGAECTNW--NSSALAQKPYSGRRPDALRLGNGHNYCRNPDR 184

QY 59 PWCYVSGEAGVPEKRPCEDLRCPEPTS 82
DB 60 PWCYVTD---PEKRYDYCDILECEB 81

RESULT 10
US-10-131-241-35
Sequence 35, Application US/10131241
Publication No. US20030012792A1
GENERAL INFORMATION:
APPLICANT: Holaday, John W.
TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
FILE REFERENCE: 05213-0344 43170-271565
CURRENT APPLICATION NUMBER: US/10/131,241
CURRENT FILING DATE: 2002-07-22
PRIOR APPLICATION NUMBER: US 09/413,049
PRIOR FILING DATE: 1999-10-06
PRIOR APPLICATION NUMBER: US 09/316,802
PRIOR FILING DATE: 1999-05-21
PRIOR APPLICATION NUMBER: US 60/086,586
PRIOR FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 65

SOFTWARE: PatentIn version 3.1
SEQ ID NO 35
LENGTH: 160
TYPE: PRT
ORGANISM: Homo sapiens
US-10-131-241-35

Query Match 11.0%; Score 142; DB 9; Length 160;
Best Local Similarity 37.2%; Pred. No. 1.3e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

Qy 4 CFMDNGHLYREDQTSAPAGRLCLNWLDA-----QSGLASAPVSGAGNHSYCRNPDEDPRG 58
Db 1 CKTGNGKXNYRGTMSTKNGITCKQWSTSPHPRFSPATHSSEGL-BENYCRNPDPNDQGS 59
Qy 59 PWCYVSGEAGVPEKR--PCEDLRCP 82
Db 60 PWCYTTD---PEKRYDYCDLLECEER 81

RESULT 11
US-09-761-120-35
Sequence 35, Application US/09761120
Patent No. US20020037847A1
GENERAL INFORMATION:
APPLICANT: Folkman, M. Judah
TITLE OF INVENTION: Nucleic Acids Encoding K-tingle 1-5 Region Fragments of Plasminogen
FILE REFERENCE: 05940-0151 (43171-252068)
CURRENT APPLICATION NUMBER: US/09/761,120
CURRENT FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 09/309,821
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 08/866,735
PRIOR FILING DATE: 1997-05-30
NUMBER OF SEQ ID NOS: 47
SOFTWARE: PatentIn version 3.0
SEQ ID NO 35
LENGTH: 160
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: K-tingle 1-2
US-09-761-120-35

Query Match 11.0%; Score 142; DB 10; Length 160;
Best Local Similarity 37.2%; Pred. No. 1.3e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

Qy 4 CFMDNGHLYREDQTSAPAGRLCLNWLDA-----QSGLASAPVSGAGNHSYCRNPDEDPRG 58
Db 1 CKTGNGKXNYRGTMSTKNGITCKQWSTSPHPRFSPATHSSEGL-BENYCRNPDPNDQGS 59
Qy 59 PWCYVSGEAGVPEKR--PCEDLRCP 82
Db 60 PWCYTTD---PEKRYDYCDLLECEER 81

RESULT 12
US-09-335-325-30
Sequence 30, Application US/09335325
Patent No. US20020164717A1
GENERAL INFORMATION:
APPLICANT: Folkman, M. Judah
O'Reilly, Michael
Cao, Yihai
Sim, B. Kim Lee
TITLE OF INVENTION: Angiostatin Fragments and Method of Use
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jones & Askew
STREET: 191 Peachtree Street, 37th Floor

CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.
ZIP: 30303-1769
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/335,325
FILING DATE: 17-Jun-1999
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/612,788
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Warren, William U.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 05213-0126
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 250 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEetical: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: K1-3
SEQUENCE DESCRIPTION: SEQ ID NO: 30:
US-09-335-325-30

Query Match 11.0%; Score 142; DB 9; Length 250;
Best Local Similarity 37.2%; Pred. No. 2.3e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

Qy 4 CFMDNGHLYREDQTSAPAGRLCLNWLDA-----QSGLASAPVSGAGNHSYCRNPDEDPRG 58
Db 1 CKTGNGKXNYRGTMSTKNGITCKQWSTSPHPRFSPATHSSEGL-BENYCRNPDPNDQGS 59
Qy 59 PWCYVSGEAGVPEKR--PCEDLRCP 82
Db 60 PWCYTTD---PEKRYDYCDLLECEER 81

RESULT 13
US-10-131-241-30
Sequence 30, Application US/10131241
Publication No. US20030012792A1
GENERAL INFORMATION:
APPLICANT: Holaday, John W.
TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
FILE REFERENCE: 05213-0344 43170-271565
CURRENT APPLICATION NUMBER: US/10/131,241
CURRENT FILING DATE: 2002-07-22
PRIOR APPLICATION NUMBER: US 09/413,049
PRIOR FILING DATE: 1999-10-06
PRIOR APPLICATION NUMBER: US 09/316,802
PRIOR FILING DATE: 1999-05-21
PRIOR APPLICATION NUMBER: US 60/086,586
PRIOR FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.1

Mon Apr 7 10:22:01 2003

us-10-057-951-2_copy_22_263.rapb

Page 6

SEQ ID NO 30
LENGTH: 250
TYPE: PRT
ORGANISM: Homo sapiens
US-10-131-241-30

Query Match 11.0%; Score 142; DB 9; Length 250;
Best Local Similarity 37.2%; Pred. No. 2.3e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

Qy 4 CFWDNGHLYREDQTSAPAGLRCLNWLDL-----QGLASAPVSGAGNHSYCRNPDEDPG 58
Db 1 CKTGNGKRYRGTMSTKXGJTCOKMSSTSPHRPSPATHPSEGL-ENYCRNPDPDPOG 59
Qy 59 PWCYVSGEAGVPEKR--PCEDLRCE 82
Db 60 PWCYTTD---PEKRYDCDILECEE 81

RESULT 14

US-09-761-120-30
Sequence 30 Application US/09761120
Patent No. US20020037847A1
GENERAL INFORMATION:
APPLICANT: Folkmann, M. Judah
APPLICANT: O'Reilly, Michael
TITLE OF INVENTION: Nucleic Acid Encoding Kringle 1-5 Region Fragments of Plasminogen
FILE REFERENCE: 05940-0151 (43171-252088)
CURRENT APPLICATION NUMBER: US/09/761,120
CURRENT FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 09/309,821
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 08/866,735
PRIOR FILING DATE: 1997-05-30
NUMBER OF SEQ ID NOS: 47
SOFTWARE: PatentIn version 3.0
SEQ ID NO 30
LENGTH: 250
TYPE: PRT
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Kringle 1-3
US-09-761-120-30

Query Match 11.0%; Score 142; DB 10; Length 250;
Best Local Similarity 37.2%; Pred. No. 2.3e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

Qy 4 CFWDNGHLYREDQTSAPAGLRCLNWLDL-----QGLASAPVSGAGNHSYCRNPDEDPG 58
Db 1 CKTGNGKRYRGTMSTKXGJTCOKMSSTSPHRPSPATHPSEGL-ENYCRNPDPDPOG 59
Qy 59 PWCYVSGEAGVPEKR--PCEDLRCE 82
Db 60 PWCYTTD---PEKRYDCDILECEE 81

RESULT 15

US-10-131-241-61
Sequence 61 Application US/10131241
Publication No. US20030012792A1
GENERAL INFORMATION:
APPLICANT: Holaday, John W.
APPLICANT: Fortier, Anne B.
TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
FILE REFERENCE: 05213-0344 43170-271565
CURRENT APPLICATION NUMBER: US/10/131,241
CURRENT FILING DATE: 2002-07-22
PRIOR APPLICATION NUMBER: US 09/413,049
PRIOR FILING DATE: 1999-10-06
PRIOR APPLICATION NUMBER: US 09/316,802

PRIOR FILING DATE: 1999-05-21
PRIOR APPLICATION NUMBER: US 60/086,586
PRIOR FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.1
SEQ ID NO 61
LENGTH: 260
TYPE: PRT
ORGANISM: Homo sapiens
US-10-131-241-61

Query Match 11.0%; Score 142; DB 9; Length 260;
Best Local Similarity 37.2%; Pred. No. 2.4e-05;
Matches 32; Conservative 8; Mismatches 34; Indels 12; Gaps 4;

Qy 4 CFWDNGHLYREDQTSAPAGLRCLNWLDL-----QGLASAPVSGAGNHSYCRNPDEDPG 58
Db 6 CKTGNGKRYRGTMSTKXGJTCOKMSSTSPHRPSPATHPSEGL-ENYCRNPDPDPOG 64
Qy 59 PWCYVSGEAGVPEKR--PCEDLRCE 82
Db 65 PWCYTTD---PEKRYDCDILECEE 86

Search completed: April 7, 2003, 09:31:28
Job time : 18.106 secs

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:17:08 ; Search time 206.782 Seconds
(without alignments)
754,542 Million cell updates/sec

Title: US-10-057-951-2_COPY_22_263

Perfect score: 1289
Sequence: 1 SGGCFMDNGHLVREDQTSFA.....PVDPGSGTFLMGQAGTREA 242

Scoring table: BLASTUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents AA Main:*

1: /cgn2_6/ptodata/1/paa/PCTUS.COMB.pep:*

2: /cgn2_6/ptodata/1/paa/US06.COMB.pep:*

3: /cgn2_6/ptodata/1/paa/US07.COMB.pep:*

4: /cgn2_6/ptodata/1/paa/US08.COMB.pep:*

5: /cgn2_6/ptodata/1/paa/US081.COMB.pep:*

6: /cgn2_6/ptodata/1/paa/US082.COMB.pep:*

7: /cgn2_6/ptodata/1/paa/US083.COMB.pep:*

8: /cgn2_6/ptodata/1/paa/US084.COMB.pep:*

9: /cgn2_6/ptodata/1/paa/US085.COMB.pep:*

10: /cgn2_6/ptodata/1/paa/US086.COMB.pep:*

11: /cgn2_6/ptodata/1/paa/US087.COMB.pep:*

12: /cgn2_6/ptodata/1/paa/US088.COMB.pep:*

13: /cgn2_6/ptodata/1/paa/US089.COMB.pep:*

14: /cgn2_6/ptodata/1/paa/US090.COMB.pep:*

15: /cgn2_6/ptodata/1/paa/US091.COMB.pep:*

16: /cgn2_6/ptodata/1/paa/US092.COMB.pep:*

17: /cgn2_6/ptodata/1/paa/US093.COMB.pep:*

18: /cgn2_6/ptodata/1/paa/US094.COMB.pep:*

19: /cgn2_6/ptodata/1/paa/US095.COMB.pep:*

20: /cgn2_6/ptodata/1/paa/US096.COMB.pep:*

21: /cgn2_6/ptodata/1/paa/US097.COMB.pep:*

22: /cgn2_6/ptodata/1/paa/US098.COMB.pep:*

23: /cgn2_6/ptodata/1/paa/US099.COMB.pep:*

24: /cgn2_6/ptodata/1/paa/US100.COMB.pep:*

25: /cgn2_6/ptodata/1/paa/US101.COMB.pep:*

26: /cgn2_6/ptodata/1/paa/US102.COMB.pep:*

27: /cgn2_6/ptodata/1/paa/US60.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1289	100.0	250 14 US-09-001-403-64	Sequence 64, Appl
2	1289	100.0	263 1 PCT-US00-27239-2	Sequence 2, Appl
3	1289	100.0	263 1 PCT-US02-05301-164	Sequence 164, App
4	1289	100.0	263 1 PCT-US98-10728-2	Sequence 2, Appl
5	1289	100.0	263 14 US-09-084-491A-2	Sequence 2, Appl
6	1289	100.0	263 15 US-09-144-889-2	Sequence 2, Appl

7	1289	100.0	263 24 US-10-057-951-2	Sequence 2, Appl
8	1289	100.0	263 25 US-10-102-704-2	Sequence 2, Appl
9	1289	100.0	263 21 US-60-048-000-2	Sequence 2, Appl
10	1286	99.8	263 27 US-09-791-537-39562	Sequence 39562, A
11	1286	99.8	263 23 US-09-927-796-44	Sequence 44, Appl
12	1286	99.8	263 26 US-10-210-951-44	Sequence 44, Appl
13	1286	99.8	263 26 US-10-211-858-44	Sequence 44, Appl
14	1286	99.8	263 26 US-10-211-884-44	Sequence 44, Appl
15	1286	99.8	263 27 US-60-230-435-1580	Sequence 1580, App
16	1286	99.8	263 27 US-60-389-987-244	Sequence 244, App
17	1286	99.8	263 27 US-60-412-418-244	Sequence 244, App
18	1275	98.9	263 20 US-09-611-526-3727	Sequence 3727, Ap
19	1235	95.8	257 1 PCT-US01-14837-8898	Sequence 8898, Ap
20	1221.5	94.8	255 1 PCT-US01-14837-8900	Sequence 8900, Ap
21	1166.5	90.5	308 27 US-60-207-315-471	Sequence 471, App
22	1063.5	82.5	286 15 US-09-144-889-4	Sequence 4, Appl
23	849	65.9	178 27 US-60-213-800-237	Sequence 237, App
24	648	50.3	146 1 PCT-US02-05301-237	Sequence 237, App
25	648	50.3	146 1 PCT-US02-05301-311	Sequence 311, App
26	596	46.2	126 17 US-09-307-140-979	Sequence 979, App
27	596	46.2	126 22 US-09-817-076-979	Sequence 979, App
28	508	39.4	109 22 US-09-834-366-13839	Sequence 13839, A
29	476	36.9	117 1 PCT-US01-14837-8897	Sequence 8897, Ap
30	426	33.0	88 22 US-09-834-366-16200	Sequence 16200, A
31	426	33.0	88 27 US-60-197-873-16200	Sequence 16200, A
32	426	33.0	88 27 US-60-197-873-16200	Sequence 16200, A
33	322	25.0	66 1 PCT-US01-00663-31583	Sequence 31583, A
34	322	25.0	66 22 US-09-864-761-38457	Sequence 38457, A
35	322	25.0	66 25 US-10-182-993-30642	Sequence 30642, A
36	322	25.0	66 25 US-10-182-995-24929	Sequence 24929, A
37	322	25.0	66 25 US-10-182-997-23626	Sequence 23626, A
38	322	25.0	66 26 US-10-203-114-31343	Sequence 31343, A
39	322	25.0	66 26 US-10-203-135-30493	Sequence 30493, A
40	322	25.0	66 26 US-10-203-136-11400	Sequence 31400, A
41	322	25.0	66 26 US-10-203-137-31583	Sequence 31583, A
42	322	25.0	66 26 US-10-203-139-30540	Sequence 30540, A
43	202	15.7	56 13 US-08-905-051-280	Sequence 238, App
44	202	15.7	56 19 US-09-547-539C-2338	Sequence 428, App
45	201	15.6	55 13 US-08-905-135-428	

ALIGNMENTS

RESULT 1

US-09-001-403-64

Sequence 64, Application US/09001403

GENERAL INFORMATION:

APPLICANT: Lal, Preeti

APPLICANT: Bardman, Olga

APPLICANT: Hillman, Jennifer L.

APPLICANT: Au-Young, Janice

APPLICANT: Tang, Y. Tom

APPLICANT: Yee, Henry

APPLICANT: Shah, Purvi

APPLICANT: Guegler, Karl J.

APPLICANT: Corley, Neil C.

TITLE OF INVENTION: HUMAN REGULATORY PROTEINS

NUMBER OF SEQUENCES: 150

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US-09/001,403

FILING DATE: HERREWTH
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: BILLINGS, LUCY J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0455 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 250 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LINDNOT05
CLONE: 3122252
US-09-001-403-64

Query Match 100.0%; Score 1289; DB 14; Length 250;
Best Local Similarity 100.0%; Pred. No. 1.1e-118;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGGCFWMDNGHLYREDQTSAPGLRCINWLDASGLASAPVSGAGNHSYCRNPDEDPKGPW 60
DB 9 SGGCFWMDNGHLYREDQTSAPGLRCINWLDASGLASAPVSGAGNHSYCRNPDEDPKGPW 68
QY 61 CYSGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIOEASGEGADEVQVPAANALPARS 120
DB 69 CYSGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIOEASGEGADEVQVPAANALPARS 128
QY 121 EAAAVQPIVIGISQVRNMSKEKDLGTLGYVLTMMVITIIAGGILLGYSYKRGDKL 180
DB 129 EAAAVQPIVIGISQVRNMSKEKDLGTLGYVLTMMVITIIAGGILLGYSYKRGDKL 188
QY 181 ECHDQVCEREMQRTLLPSAFTNPCEIYDEKTVVHTSQTVPVPOEGSTPLMGOAGTP 240
DB 189 ECHDQVCEREMQRTLLPSAFTNPCEIYDEKTVVHTSQTVPVPOEGSTPLMGOAGTP 248
QY 241 GA 242
DB 249 GA 250

RESULT 2
PCT-US00-27239-2
Sequence 2, Application PC/TUS0027239
GENERAL INFORMATION:
APPLICANT: HUMAN GENOME SCIENCES, INC.
TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
FILE REFERENCE: PF378PCT2
CURRENT APPLICATION NUMBER: PCT/US00/27239
CURRENT FILING DATE: 2000-10-03
PRIOR FILING DATE: 09/41,977
PRIOR FILING DATE: 1999-10-04
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US00-27239-2

Query Match 100.0%; Score 1289; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.2e-118;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SGGCFWMDNGHLYREDQTSAPGLRCINWLDASGLASAPVSGAGNHSYCRNPDEDPKGPW 60
DB 22 SGGCFWMDNGHLYREDQTSAPGLRCINWLDASGLASAPVSGAGNHSYCRNPDEDPKGPW 81
QY 61 CYSGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIOEASGEGADEVQVPAANALPARS 120

DB 82 CYSGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIOEASGEGADEVQVPAANALPARS 141
QY 121 EAAAVQPIVIGISQVRNMSKEKDLGTLGYVLTMMVITIIAGGILLGYSYKRGDKL 180
DB 142 EAAAVQPIVIGISQVRNMSKEKDLGTLGYVLTMMVITIIAGGILLGYSYKRGDKL 201
QY 181 ECHDQVCEREMQRTLLPSAFTNPCEIYDEKTVVHTSQTVPVPOEGSTPLMGOAGTP 240
DB 202 ECHDQVCEREMQRTLLPSAFTNPCEIYDEKTVVHTSQTVPVPOEGSTPLMGOAGTP 261
QY 241 GA 242
DB 262 GA 263

RESULT 3
PCT-US02-05301-164
Sequence 164, Application PC/TUS0205301
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: 70 Human Secreted Proteins
FILE REFERENCE: PS736PCT
CURRENT APPLICATION NUMBER: PCT/US02/05301
CURRENT FILING DATE: 2002-02-21
PRIOR APPLICATION NUMBER: US 60/304,417
PRIOR FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: US 60/270,625
PRIOR FILING DATE: 2001-02-23
NUMBER OF SEQ ID NOS: 340
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 164
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US02-05301-164

Query Match 100.0%; Score 1289; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.2e-118;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGGCFWMDNGHLYREDQTSAPGLRCINWLDASGLASAPVSGAGNHSYCRNPDEDPKGPW 60
DB 22 SGGCFWMDNGHLYREDQTSAPGLRCINWLDASGLASAPVSGAGNHSYCRNPDEDPKGPW 81
QY 61 CYSGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIOEASGEGADEVQVPAANALPARS 120
DB 82 CYSGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIOEASGEGADEVQVPAANALPARS 141
QY 121 EAAAVQPIVIGISQVRNMSKEKDLGTLGYVLTMMVITIIAGGILLGYSYKRGDKL 180
DB 142 EAAAVQPIVIGISQVRNMSKEKDLGTLGYVLTMMVITIIAGGILLGYSYKRGDKL 201
QY 181 ECHDQVCEREMQRTLLPSAFTNPCEIYDEKTVVHTSQTVPVPOEGSTPLMGOAGTP 240
DB 202 ECHDQVCEREMQRTLLPSAFTNPCEIYDEKTVVHTSQTVPVPOEGSTPLMGOAGTP 261
QY 241 GA 242
DB 262 GA 263

RESULT 4
PCT-US98-10728-2
Sequence 2, Application PC/TUS9810728
GENERAL INFORMATION:
APPLICANT: EBNER, REINHARD
APPLICANT: MOORE, PAUL
APPLICANT: RUBEN, STEVEN
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.


```

; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/10728
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A.
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF379PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8439
; TELEFAX: (301) 309-8504
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US98-10728-2

Query Match          100.0%; Score 1289; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.2e-118;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGGCFWMDNGLYREDQTSPPAGRLCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGFW 60
DB 22 SGGCFWMDNGLYREDQTSPPAGRLCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGFW 81
QY 61 CVYSGAGVPEKRPCEDLRCPEITTSQALPAFTTEIOEASGPGADEVOYFAPANALPARS 120
DB 82 CVYSGAGVPEKRPCEDLRCPEITTSQALPAFTTEIOEASGPGADEVOYFAPANALPARS 141
QY 121 EAAAVOPVIGISORVRMNSKEKDDGLTGLVGLITMWVILIAIGAGIILIGSYKRGKDLK 180
DB 142 EAAAVOPVIGISORVRMNSKEKDDGLTGLVGLITMWVILIAIGAGIILIGSYKRGKDLK 201
QY 181 EOHDKVCEREMORITLPLSAFTNPTCEIYDEKTVVHTSQTPVDPQEGSTPLMGQAGTP 240
DB 202 EOHDKVCEREMORITLPLSAFTNPTCEIYDEKTVVHTSQTPVDPQEGSTPLMGQAGTP 261
QY 241 GA 242
DB 262 GA 263

RESULT 5
US-09-084-491A-2
; Sequence 2, Application US/09084491A
; GENERAL INFORMATION:
; APPLICANT: MOORE, PAUL A.
; APPLICANT: RUBEN, STEVEN M.
; APPLICANT: RUBEN, REINHARD
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/10728
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A.
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF379PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8439
; TELEFAX: (301) 309-8504
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US98-10728-2
```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/084,491A
; FILING DATE: 27-MAY-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A.
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF378
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-084-491A-2

Query Match          100.0%; Score 1289; DB 14; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.2e-118;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGGCFWMDNGLYREDQTSPPAGRLCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGFW 60
DB 22 SGGCFWMDNGLYREDQTSPPAGRLCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGFW 81
QY 61 CVYSGAGVPEKRPCEDLRCPEITTSQALPAFTTEIOEASGPGADEVOYFAPANALPARS 120
DB 82 CVYSGAGVPEKRPCEDLRCPEITTSQALPAFTTEIOEASGPGADEVOYFAPANALPARS 141
QY 121 EAAAVOPVIGISORVRMNSKEKDDGLTGLVGLITMWVILIAIGAGIILIGSYKRGKDLK 180
DB 142 EAAAVOPVIGISORVRMNSKEKDDGLTGLVGLITMWVILIAIGAGIILIGSYKRGKDLK 201
QY 181 EOHDKVCEREMORITLPLSAFTNPTCEIYDEKTVVHTSQTPVDPQEGSTPLMGQAGTP 240
DB 202 EOHDKVCEREMORITLPLSAFTNPTCEIYDEKTVVHTSQTPVDPQEGSTPLMGQAGTP 261
QY 241 GA 242
DB 262 GA 263

RESULT 6
US-09-144-889-2
; Sequence 2, Application US/09144889B
; GENERAL INFORMATION:
; APPLICANT: EARL F. ALBONE
; APPLICANT: KRISTINE K. KIKLY
; TITLE OF INVENTION: KRINGLE1
; FILE REFERENCE: GH-70249
; CURRENT APPLICATION NUMBER: US/09/144,889B
; CURRENT FILING DATE: 1998-09-01
; EARLIER APPLICATION NUMBER: 60/056,032
; EARLIER FILING DATE: 1997-09-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
; US-09-144-889-2

Query Match          100.0%; Score 1289; DB 15; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.2e-118;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGGCFWMDNGLYREDQTSPPAGRLCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGFW 60
DB 22 SGGCFWMDNGLYREDQTSPPAGRLCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGFW 81
```

```
QY 61 CVYSGAGVPEKRPCEDELRCPETTSOALPAFTTEIOEASBGADVOYFAPANALPARS 120
| 82 CVYSGAGVPEKRPCEDELRCPETTSOALPAFTTEIOEASBGADVOYFAPANALPARS 141
|
QY 121 EAAAVQPVIGISORVRNMSKEKDLGTLGYVLGITMNVIIAIGAGIILGYSKRGKDLK 180
| 142 EAAAVQPVIGISORVRNMSKEKDLGTLGYVLGITMNVIIAIGAGIILGYSKRGKDLK 201
|
QY 181 EOHDOVCEREMQRTITLPLSAFTNPCEIYDEKTVVHTSQTVPDQEGSTPLMGAGTTP 240
| 202 EOHDOVCEREMQRTITLPLSAFTNPCEIYDEKTVVHTSQTVPDQEGSTPLMGAGTTP 261
|
QY 241 GA 242
| 262 GA 263
```

```
RESULT 7
US-10-057-951-2
; Sequence 2, Application US/10057951
; GENERAL INFORMATION:
; APPLICANT: Moore et al.
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
; FILE REFERENCE: PF378P1
; CURRENT APPLICATION NUMBER: US/10/057,951
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US 09/411,977
; PRIOR FILING DATE: 1999-10-04
; PRIOR APPLICATION NUMBER: US 09/084,491
; PRIOR FILING DATE: 1998-05-27
; PRIOR APPLICATION NUMBER: US 60/048,000
; PRIOR FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-951-2
```

```
Query Match 100.0%; Score 1289; DB 24; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.2e-118;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 SGGCFWMDNGHLTYREDQTSAPGLRCLNWLDAQSLASAPVSGAGNHSYCRNPDEDPGRGPW 60
| 22 SGGCFWMDNGHLTYREDQTSAPGLRCLNWLDAQSLASAPVSGAGNHSYCRNPDEDPGRGPW 81
|
QY 61 CVYSGAGVPEKRPCEDELRCPETTSOALPAFTTEIOEASBGADVOYFAPANALPARS 120
| 82 CVYSGAGVPEKRPCEDELRCPETTSOALPAFTTEIOEASBGADVOYFAPANALPARS 141
|
QY 121 EAAAVQPVIGISORVRNMSKEKDLGTLGYVLGITMNVIIAIGAGIILGYSKRGKDLK 180
| 142 EAAAVQPVIGISORVRNMSKEKDLGTLGYVLGITMNVIIAIGAGIILGYSKRGKDLK 201
|
QY 181 EOHDOVCEREMQRTITLPLSAFTNPCEIYDEKTVVHTSQTVPDQEGSTPLMGAGTTP 240
| 202 EOHDOVCEREMQRTITLPLSAFTNPCEIYDEKTVVHTSQTVPDQEGSTPLMGAGTTP 261
|
QY 241 GA 242
| 262 GA 263
```

```
RESULT 8
US-10-102-704-2
; Sequence 2, Application US/10102704
; GENERAL INFORMATION:
; APPLICANT: Moore et al.
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protein
; FILE REFERENCE: PF378CI
```

```
; CURRENT APPLICATION NUMBER: US/10/102,704
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/084,491
; PRIOR FILING DATE: 1998-05-27
; PRIOR APPLICATION NUMBER: 60/048,000
; PRIOR FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-102-704-2
```

```
Query Match 100.0%; Score 1289; DB 25; Length 263;
Best Local Similarity 100.0%; Pred No. 1.2e-118;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 SGGCFWMDNGHLTYREDQTSAPGLRCLNWLDAQSLASAPVSGAGNHSYCRNPDEDPGRGPW 60
| 22 SGGCFWMDNGHLTYREDQTSAPGLRCLNWLDAQSLASAPVSGAGNHSYCRNPDEDPGRGPW 81
|
QY 61 CVYSGAGVPEKRPCEDELRCPETTSOALPAFTTEIOEASBGADVOYFAPANALPARS 120
| 82 CVYSGAGVPEKRPCEDELRCPETTSOALPAFTTEIOEASBGADVOYFAPANALPARS 141
|
QY 121 EAAAVQPVIGISORVRNMSKEKDLGTLGYVLGITMNVIIAIGAGIILGYSKRGKDLK 180
| 142 EAAAVQPVIGISORVRNMSKEKDLGTLGYVLGITMNVIIAIGAGIILGYSKRGKDLK 201
|
QY 181 EOHDOVCEREMQRTITLPLSAFTNPCEIYDEKTVVHTSQTVPDQEGSTPLMGAGTTP 240
| 202 EOHDOVCEREMQRTITLPLSAFTNPCEIYDEKTVVHTSQTVPDQEGSTPLMGAGTTP 261
|
QY 241 GA 242
| 262 GA 263
```

```
RESULT 9
US-60-048-000-2
; Sequence 2, Application US/60048000
; GENERAL INFORMATION:
; APPLICANT: EBNER, REINHARD
; APPLICANT: MOORE, PAUL
; APPLICANT: RUBEN, STEVE
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSER: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/048,000
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A.
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF379PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 2:
```

SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-60-048-000-2

Query Match 100.0%; Score 1289; DB 21; Length 263;
Best Local Similarity 100.0%; Pred. No. 1,2e-118;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGGCFMNGHLVREDQTSAPGLRCINWLDAGSLASAPVSGAGNHSYCRNPDDEPRGPW 60
DB 22 SGGCFMNGHLVREDQTSAPGLRCINWLDAGSLASAPVSGAGNHSYCRNPDDEPRGPW 81
QY 61 CYSSEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASBEGADDEVQFAPANALPARS 120
DB 82 CYSSEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASBEGADDEVQFAPANALPARS 141
QY 121 EAAAVQPVIGISQVRMNSKEKKDLGTLGYVIGITMVIILAGIILGYSYKRGDKL 180
DB 142 EAAAVQPVIGISQVRMNSKEKKDLGTLGYVIGITMVIILAGIILGYSYKRGDKL 201
QY 181 EOHDKVCEREMORTLPLSAFTNPCEIYDEKTVVHTSQFPVDPQEGSTPLMGQAGTP 240
DB 202 EOHDKVCEREMORTLPLSAFTNPCEIYDEKTVVHTSQFPVDPQEGSTPLMGQAGTP 261
QY 241 GA 242
DB 262 GA 263

RESULT 10
US-09-791-537-39562
Sequence 39562, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
APPLICANT: Dancer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patent version 3.0
SEQ ID NO 39562
LENGTH: 263
TYPE: PRP
ORGANISM: Homo sapiens
US-09-791-537-39562

Query Match 99.8%; Score 1286; DB 21; Length 263;
Best Local Similarity 99.6%; Pred. No. 2,3e-118;
Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGGCFMNGHLVREDQTSAPGLRCINWLDAGSLASAPVSGAGNHSYCRNPDDEPRGPW 60
DB 22 SGGCFMNGHLVREDQTSAPGLRCINWLDAGSLASAPVSGAGNHSYCRNPDDEPRGPW 81
QY 61 CYSSEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASBEGADDEVQFAPANALPARS 120
DB 82 CYSSEAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASBEGADDEVQFAPANALPARS 141
QY 121 EAAAVQPVIGISQVRMNSKEKKDLGTLGYVIGITMVIILAGIILGYSYKRGDKL 180
DB 142 EAAAVQPVIGISQVRMNSKEKKDLGTLGYVIGITMVIILAGIILGYSYKRGDKL 201
QY 181 EOHDKVCEREMORTLPLSAFTNPCEIYDEKTVVHTSQFPVDPQEGSTPLMGQAGTP 240
DB 202 EOHDKVCEREMORTLPLSAFTNPCEIYDEKTVVHTSQFPVDPQEGSTPLMGQAGTP 261
QY 241 GA 242

DB 262 GA 263

RESULT 11
US-09-927-796-44
Sequence 44, Application US/09927796
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Goddard, Audrey
APPLICANT: Goddard, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Marsters, Scot A.
APPLICANT: Pan, James
APPLICANT: Pittl, Robert M.
APPLICANT: Roy, Margaret Ann
APPLICANT: Smith, Victoria
APPLICANT: Stone, Donna M.
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
FILE REFERENCE: P2931R1C1
CURRENT APPLICATION NUMBER: US/09/927,796
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 60/014699
PRIOR FILING DATE: 1996-04-01
PRIOR APPLICATION NUMBER: 60/026943
PRIOR FILING DATE: 1996-09-23
PRIOR APPLICATION NUMBER: 60/059121
PRIOR FILING DATE: 1997-07-17
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/062037
PRIOR FILING DATE: 1997-10-10
PRIOR APPLICATION NUMBER: 60/063755
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063045
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063046
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/066511
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066772
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/067411
PRIOR FILING DATE: 1997-12-03
PRIOR APPLICATION NUMBER: 60/069862
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/095929
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/097978
PRIOR FILING DATE: 1998-08-26
PRIOR APPLICATION NUMBER: 60/103396
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: 60/108867
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/112851
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/119965
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: 60/123972
PRIOR FILING DATE: 1999-03-11
PRIOR APPLICATION NUMBER: 60/133459
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 60/140650
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/140653
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/144758
PRIOR FILING DATE: 1999-07-20

```

PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/146222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: 60/149395
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/151689
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 08/625328
PRIOR FILING DATE: 1996-04-01
PRIOR APPLICATION NUMBER: 08/710802
PRIOR FILING DATE: 1996-09-23
PRIOR APPLICATION NUMBER: 08/800699
PRIOR FILING DATE: 1997-02-14
PRIOR APPLICATION NUMBER: 08/828683
PRIOR FILING DATE: 1997-03-31
PRIOR APPLICATION NUMBER: 08/829270
PRIOR FILING DATE: 1997-03-31
PRIOR APPLICATION NUMBER: 08/928069
PRIOR FILING DATE: 1997-09-11
PRIOR APPLICATION NUMBER: 08/934494
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 09/143068
PRIOR FILING DATE: 1998-08-28
PRIOR APPLICATION NUMBER: 09/143707
PRIOR FILING DATE: 1998-08-28
PRIOR APPLICATION NUMBER: 09/151889
PRIOR FILING DATE: 1998-09-11
PRIOR APPLICATION NUMBER: 09/169104
PRIOR FILING DATE: 1998-10-09
PRIOR APPLICATION NUMBER: 09/202089
PRIOR FILING DATE: 1998-12-08
PRIOR APPLICATION NUMBER: 09/254311
PRIOR FILING DATE: 1999-03-03
PRIOR APPLICATION NUMBER: 09/304003
PRIOR FILING DATE: 1999-04-30
PRIOR APPLICATION NUMBER: 09/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380138
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380139
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/403297
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/423844
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 09/511133
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 09/511631
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 09/646410
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: 09/665350
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: 09/690169
PRIOR FILING DATE: 2000-10-16
PRIOR APPLICATION NUMBER: 09/690189
PRIOR FILING DATE: 2000-10-16
PRIOR APPLICATION NUMBER: 09/709238
PRIOR FILING DATE: 2000-11-18
PRIOR APPLICATION NUMBER: 09/866034
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 09/872035
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 09/884733
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: 09/886342
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: 09/866028
PRIOR FILING DATE: 2001-08-25
PRIOR APPLICATION NUMBER: PCT/US97/05230
PRIOR FILING DATE: 1997-03-31
PRIOR APPLICATION NUMBER: PCT/US98/19094
```

```

PRIOR FILING DATE: 1998-09-14
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: PCT/US98/21407
PRIOR FILING DATE: 1998-10-09
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: 1998-12-01
PRIOR APPLICATION NUMBER: PCT/US99/05028
PRIOR FILING DATE: 1999-03-08
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: PCT/US99/20111
PRIOR FILING DATE: 1999-09-01
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28301
PRIOR FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: PCT/US99/28634
PRIOR FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: PCT/US00/04341
PRIOR FILING DATE: 2000-02-18
PRIOR APPLICATION NUMBER: PCT/US00/04342
PRIOR FILING DATE: 2000-02-18
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/06884
PRIOR FILING DATE: 2000-03-15
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: PCT/US00/13705
PRIOR FILING DATE: 2000-05-17
PRIOR APPLICATION NUMBER: PCT/US00/14941
PRIOR FILING DATE: 2000-05-30
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/20710
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: PCT/US01/17800
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: PCT/US01/19692
PRIOR FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: PCT/US01/21066
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: PCT/US01/21735
PRIOR FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 258
SEQ ID NO 44
LENGTH: 263
```

```

Query Match          99.8%; Score 1286; DB 23; Length 263;
Best Local Similarity 99.6%; Pred. No. 2.3e-118;
Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 SGGCFWNGHLYREDQSPAPGRCLNMLDAQGLASAPYSGAGNHSYCNPDDEDRGPW 60
Db 22 SGGCFWNGHLYREDQSPAPGRCLNMLDAQGLASAPYSGAGNHSYCNPDDEDRGPW 81
QY 61 CYSVGEAVDEKRCPCEDLRCPETTSQALPAFTTEIOBASGPADEVQVAPANALPARS 120
Db 82 CYSVGEAVDEKRCPCEDLRCPETTSQALPAFTTEIOBASGPADEVQVAPANALPARS 141
```

```
QY 121 EAAVOPVIGISORVRMSKEKDLGTLGYVLGTTMMVITIAAGITIGSYKRGKDLK 180
DB 142 EAAVOPVIGISORVRMSKEKDLGTLGYVLGTTMMVITIAAGITIGSYKRGKDLK 201
QY 181 EOHDKVCEREMQRTITLPLSAFTNPTCEIYDEKTVVHTSQTPVDPQESGTPLMQAGTP 240
DB 202 EOHDKVCEREMQRTITLPLSAFTNPTCEIYDEKTVVHTSQTPVDPQESGTPLMQAGTP 261
QY 241 GA 242
DB 262 GA 263
```

RESULT 12

```
US-10-210-951-44
; Sequence 44, Application US/10210951
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Masters, Scott A.
; APPLICANT: Pan, James
; APPLICANT: Pitti, Robert M.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stone, Donna M.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
; FILE REFERENCE: P2931R1C1
; CURRENT APPLICATION NUMBER: US/10/210,951
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/014699
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 60/026943
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 60/059121
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/062037
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 258
; SEQ ID NO 44
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-210-951-44
```

```
Query Match 99.8%; Score 1286; DB 26; Length 263;
Best Local Similarity 99.6%; Pred. No. 2,3e-118;
Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SGGCFMNDNHLIYREDOTSPAPGLRCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDRGPW 60
DB 22 SGGCFMNDNHLIYREDOTSPAPGLRCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDRGPW 81
QY 61 CYYSGAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASBGPADEVQVFAPANALPARS 120
DB 82 CYYSGAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASBGPADEVQVFAPANALPARS 141
```

```
QY 121 EAAVOPVIGISORVRMSKEKDLGTLGYVLGTTMMVITIAAGITIGSYKRGKDLK 180
DB 142 EAAVOPVIGISORVRMSKEKDLGTLGYVLGTTMMVITIAAGITIGSYKRGKDLK 201
QY 181 EOHDKVCEREMQRTITLPLSAFTNPTCEIYDEKTVVHTSQTPVDPQESGTPLMQAGTP 240
DB 202 EOHDKVCEREMQRTITLPLSAFTNPTCEIYDEKTVVHTSQTPVDPQESGTPLMQAGTP 261
QY 241 GA 242
DB 262 GA 263
```

RESULT 13

```
US-10-211-858-44
; Sequence 44, Application US/10211858
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Masters, Scott A.
; APPLICANT: Pan, James
; APPLICANT: Pitti, Robert M.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stone, Donna M.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
; FILE REFERENCE: P2931R1C1
; CURRENT APPLICATION NUMBER: US/10/211,858
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/014699
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 60/026943
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 60/059121
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/062037
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 258
; SEQ ID NO 44
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-858-44
```

```
Query Match 99.8%; Score 1286; DB 26; Length 263;
Best Local Similarity 99.6%; Pred. No. 2,3e-118;
Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SGGCFMNDNHLIYREDOTSPAPGLRCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDRGPW 60
DB 22 SGGCFMNDNHLIYREDOTSPAPGLRCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDRGPW 81
QY 61 CYYSGAGVPEKRPCEDLRCPETTSQALPAFTTEIOEASBGPADEVQVFAPANALPARS 120
```

```
Db      82  CYSVSGAGVPEKRPCEDLRCPEITTSQALPAFTTEIQEASGEGADEVQYFAPANALPARS 141
Qy      121  EAAAVQPVIGISQVRNMSKEKKDLGTLGYVLGITMMVIIAIGAGIILGYSYKRGKDLK 180
Db      142  EAAAVQPVIGISQVRNMSKEKKDLGTLGYVLGITMMVIIAIGAGIILGYSYKRGKDLK 201
Qy      181  EOHDKVCEREMORITLPLSAFTNPCEIYDEKTVVHTSQTPVDPQEGSTPLMGQAGTP 240
Db      202  EOHDKVCEREMORITLPLSAFTNPCEIYDEKTVVHTSQTPVDPQEGSTPLMGQAGTP 261
Qy      241  GA 242
Db      262  GA 263

RESULT 14
US-10-211-884-44
; Sequence 44, Application US/10211884
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Marsters, Scot A.
; APPLICANT: Pan, James
; APPLICANT: Pizzi, Robert M.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stone, Donna M.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
; FILE REFERENCE: P2931R1C1
; CURRENT APPLICATION NUMBER: US/10/211,884
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/014699
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 60/026943
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 60/059121
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/062037
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 258
; SEQ ID NO 44
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-884-44

Query Match      99.8%; Score 1286; DB 26; Length 263;
Best Local Similarity 99.6%; Pred. No. 2.3e-118;
Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      14  SGCCFMDNGHLTYREDQTSAPGLRCLNWLDAQSGIASAPVSGAGNHSYCRNPDEDPRGPW 60
Db      22  SGCCFMDNGHLTYREDQTSAPGLRCLNWLDAQSGIASAPVSGAGNHSYCRNPDEDPRGPW 81
Qy      61  CYSVSGAGVPEKRPCEDLRCPEITTSQALPAFTTEIQEASGEGADEVQYFAPANALPARS 120
```

```
Db      82  CYSVSGAGVPEKRPCEDLRCPEITTSQALPAFTTEIQEASGEGADEVQYFAPANALPARS 141
Qy      121  EAAAVQPVIGISQVRNMSKEKKDLGTLGYVLGITMMVIIAIGAGIILGYSYKRGKDLK 180
Db      142  EAAAVQPVIGISQVRNMSKEKKDLGTLGYVLGITMMVIIAIGAGIILGYSYKRGKDLK 201
Qy      181  EOHDKVCEREMORITLPLSAFTNPCEIYDEKTVVHTSQTPVDPQEGSTPLMGQAGTP 240
Db      202  EOHDKVCEREMORITLPLSAFTNPCEIYDEKTVVHTSQTPVDPQEGSTPLMGQAGTP 261
Qy      241  GA 242
Db      262  GA 263

RESULT 15
US-60-230-435-1580
; Sequence 1580, Application US/60230435
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: C1000768
; CURRENT APPLICATION NUMBER: US/60/230,435
; CURRENT FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 2991
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1580
; LENGTH: 263
; TYPE: PRT
; ORGANISM: HUMAN
US-60-230-435-1580
```

```
Query Match      99.8%; Score 1286; DB 27; Length 263;
Best Local Similarity 99.6%; Pred. No. 2.3e-118;
Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1  SGCCFMDNGHLTYREDQTSAPGLRCLNWLDAQSGIASAPVSGAGNHSYCRNPDEDPRGPW 60
Db      22  SGCCFMDNGHLTYREDQTSAPGLRCLNWLDAQSGIASAPVSGAGNHSYCRNPDEDPRGPW 81
Qy      61  CYSVSGAGVPEKRPCEDLRCPEITTSQALPAFTTEIQEASGEGADEVQYFAPANALPARS 120
Db      82  CYSVSGAGVPEKRPCEDLRCPEITTSQALPAFTTEIQEASGEGADEVQYFAPANALPARS 141
Qy      121  EAAAVQPVIGISQVRNMSKEKKDLGTLGYVLGITMMVIIAIGAGIILGYSYKRGKDLK 180
Db      142  EAAAVQPVIGISQVRNMSKEKKDLGTLGYVLGITMMVIIAIGAGIILGYSYKRGKDLK 201
Qy      181  EOHDKVCEREMORITLPLSAFTNPCEIYDEKTVVHTSQTPVDPQEGSTPLMGQAGTP 240
Db      202  EOHDKVCEREMORITLPLSAFTNPCEIYDEKTVVHTSQTPVDPQEGSTPLMGQAGTP 261
Qy      241  GA 242
Db      262  GA 263
```

```
Search completed: April 7, 2003, 09:29:26
Job time : 207.782 secs
```

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:18:38 ; Search time 36.2245 Seconds

(without alignments)
1010.317 Million cell updates/sec

Title: US-10-057-951-2_COPY_22_263

Perfect score: 1289

Sequence: 1 SGCCFMDNGHLHYREDQTSFA.....PVDQEGSTPLMGAGTPGA 242

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 679521 seqs, 151232488 residues

Total number of hits satisfying chosen parameters: 679521

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/prodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/prodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/prodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/prodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/prodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/prodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1280	99.3	263	6	US-10-218-140-6002
2	1028.5	79.8	264	6	US-10-144-779-488
3	154.5	12.0	293	7	US-60-452-680-14406
4	154.5	12.0	293	7	US-60-453-135-8958
5	154.5	12.0	293	7	US-60-453-050-8958
6	154.5	12.0	516	7	US-60-452-680-14405
7	154.5	12.0	516	7	US-60-453-135-8957
8	154.5	12.0	516	7	US-60-453-050-8957
9	154.5	12.0	527	5	US-09-612-314-51
10	154.5	12.0	562	1	PCT-US02-32263-26
11	154.5	12.0	562	5	US-09-703-695A-4
12	154.5	12.0	562	6	US-10-287-994-26
13	154.5	12.0	562	7	US-60-452-680-14408
14	154.5	12.0	562	7	US-60-453-135-8960
15	154.5	12.0	562	7	US-60-453-050-8960
16	154.5	12.0	587	5	US-09-949-016-11501
17	146.5	11.4	655	1	PCT-US02-19017-28
18	146.5	11.4	656	7	US-60-453-135-11516
19	146.5	11.4	656	7	US-60-453-050-11516
20	146	11.3	322	1	PCT-US02-27855-21
21	146	11.3	322	1	PCT-US02-27855A-21
22	143.5	11.1	482	7	US-60-452-680-14407
23	143.5	11.1	482	7	US-60-453-135-8959
24	143.5	11.1	482	7	US-60-453-050-8959
25	143.5	11.1	482	7	US-60-453-050-8959
26	142	11.0	339	6	US-10-401-108-3

27	142	11.0	363	6	US-10-292-418-11	Sequence 11, Appl
28	142	11.0	391	1	PCT-US02-37879-7	Sequence 7, Appl
29	142	11.0	391	6	US-10-304-287-7	Sequence 7, Appl
30	142	11.0	394	1	PCT-US02-37879-8	Sequence 8, Appl
31	142	11.0	394	6	US-10-304-287-8	Sequence 8, Appl
32	142	11.0	453	5	US-09-291-200A-1	Sequence 1, Appl
33	142	11.0	453	5	US-09-946-893B-4	Sequence 4, Appl
34	142	11.0	569	5	US-09-946-893B-5	Sequence 5, Appl
35	142	11.0	571	5	US-09-946-893B-8	Sequence 8, Appl
36	142	11.0	576	5	US-09-946-893B-6	Sequence 6, Appl
37	142	11.0	713	5	US-09-949-016-9983	Sequence 9983, Ap
38	142	11.0	713	7	US-60-452-680-22399	Sequence 22399, A
39	142	11.0	713	7	US-60-453-135-13847	Sequence 13847, A
40	142	11.0	713	7	US-60-453-050-13847	Sequence 13847, A
41	142	11.0	791	1	PCT-US02-37879-1	Sequence 1, Appl
42	142	11.0	791	6	US-10-304-287-1	Sequence 1, Appl
43	142	11.0	810	5	US-09-291-200A-5	Sequence 5, Appl
44	142	11.0	810	5	US-09-946-893B-2	Sequence 2, Appl
45	135.5	10.5	688	1	PCT-US02-27855-18	Sequence 18, Appl

ALIGNMENTS

```
RESULT 1
US-10-218-140-6002
Sequence 6002, Application US/10218140
GENERAL INFORMATION:
APPLICANT: Leach, Martin D.
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES
FILE REFERENCE: 15966-543 CON
CURRENT FILING DATE: 2002-08-12
PRIOR FILING DATE: 2002-08-12
PRIOR APPLICATION NUMBER: 09/540,763
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: 60/127,728
PRIOR FILING DATE: 1999-04-05
PRIOR APPLICATION NUMBER: 60/127,636
PRIOR FILING DATE: 1999-04-02
PRIOR APPLICATION NUMBER: 60/127,607
PRIOR FILING DATE: 1999-03-31
NUMBER OF SEQ ID NOS: 6322
SOFTWARE: Curator Version 1.0
SEQ ID NO 6002
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
US-10-218-140-6002
Query Match 99.3%; Score 1280; DB 6; Length 263;
Best Local Similarity 99.2%; Pred. No. 5.4e-102;
Matches 240; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
1 SGCCFMDNGHLHYREDQTSFAGLRCLNWLDAQGLASAPVSGAGNHSYCRNDEDPGRGW 60
|||||
22 SGCCFMDNGHLHYREDQTSFAGLRCLNWLDAQGLASAPVSGAGNHSYCRNDEDPGRGW 81
61 CVVSGEAGVPEKRPEDJRCPTTSOALPAFTTEIOEASGEPADBEVQVAPANALPARS 120
82 CVVSEAGVPEKRPEDJRCPTTSOALPAFTTEIOEASGEPADBEVQVAPANALPARS 141
121 EAAAVQPIVIGISQRRNRSKKKDLGTLGYLVLTMMVITIAIGAGIIGYGYKKGDKL 180
142 EAAAVQPIVIGISQRRNRSKKKDLGTLGYLVLTMMVITIAIGAGIIGYGYKKGDKL 201
181 EOHDKVCEKEMQRTTLPISAFNPTCEIYDEKTVVHTSQTPVDEQEGSTPLMGAGTGP 240
202 EOHDKVCEKEMQRTTLPISAFNPTCEIYDEKTVVHTSQTPVDEQEGSTPLMGAGTGP 261
241 GA 242
|||
```

Db 262 GA 263

RESULT 2
US-10-144-779-488

; Sequence 488, Application US/10144779
; GENERAL INFORMATION:
; APPLICANT: SUBRAMANIAN, Mani et al.
; TITLE OF INVENTION: MOUSE ORTHOLOGS OF HUMAN DISEASE GENES.
; FILE REFERENCE: CL001235
; CURRENT APPLICATION NUMBER: US/10/144,779
; CURRENT FILING DATE: 2002-05-15
; NUMBER OF SEQ ID NOS: 864
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 488
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-144-779-488

Query Match 79.8%; Score 1028.5; DB 6; Length 264;
Best Local Similarity 78.6%; Pred. No. 2.2e-80;
Matches 191; Conservative 16; Mismatches 35; Indels 1; Gaps 1;

QY 1 SGGCFMNDGHLVREDQTSAPAGRLCLNWLDAQSGLASAPVSGAGNHXYCRNPDEDPGRGPW 60
DB 22 SGGCFMNDGHLVREDQTSAPAGRLCLNWLDAQSGRESLTPSPGNHNYCRNPDDQPRGPW 81
QY 61 CTVSGEAGVPEKRPCEDLRCPEPTTSA-LPATTETQESSEGPADDEVQVAPANALPAR 119
DB 82 CTVSSSTGVPEKRPCEDLRCPEPTTSAQPPSSAMELEKSSGAPDKEQVPPANALPAR 141
QY 120 SEAAVQVPVIGISQRYRMSKEKDLGTGLGYLGTMMVYIIAGIITLGYSYRGKDL 179
DB 142 SEAAVQVPVIGISQRYRMSKEKDLGTGLGYLGTMMVYIIAGIITLGYSYRGKDL 201
QY 180 KECHDKVCEREMQITLLPLSAFTNPTCEIIDEKTVVHTSQTPVDPEGSTPLMGAGT 239
DB 202 KECHDKVCEREMQITLLPLSAFTNPTCEIIDEKTVVHTSQTPVDPEGSTPLMGAGT 261
QY 240 PGA 242
DB 262 PGA 264

RESULT 3
US-60-452-680-14406

; Sequence 14406, Application US/60452680
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: GROPE, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001450
; CURRENT APPLICATION NUMBER: US/60/452,680
; CURRENT FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14406
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-452-680-14406

Query Match 12.0%; Score 154.5; DB 7; Length 293;
Best Local Similarity 39.6%; Pred. No. 2.8e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 4 CFMDNGHLVREDQTSAPAGRLCLNWLDAQSGLASAPVSGAGNHXYCRNPDE 54
DB 127 CYEDGISTRGWSTAESGAECTNM--NSSALAQKPYSGRRPDALIRLGLGNHNYCRNPDR 184

QY 55 DPGFPCVYSGEAGVPEKRPCEDLRCPEPTTS 85
DB 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 213

RESULT 4
US-60-453-135-8958

; Sequence 8958, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8958
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-8958

Query Match 12.0%; Score 154.5; DB 7; Length 293;
Best Local Similarity 39.6%; Pred. No. 2.8e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 4 CFMDNGHLVREDQTSAPAGRLCLNWLDAQSGLASAPVSGAGNHXYCRNPDE 54
DB 127 CYEDGISTRGWSTAESGAECTNM--NSSALAQKPYSGRRPDALIRLGLGNHNYCRNPDR 184
QY 55 DPGFPCVYSGEAGVPEKRPCEDLRCPEPTTS 85
DB 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 213

RESULT 5
US-60-453-050-8958

; Sequence 8958, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8958
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-8958

Query Match 12.0%; Score 154.5; DB 7; Length 293;
Best Local Similarity 39.6%; Pred. No. 2.8e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 4 CFMDNGHLVREDQTSAPAGRLCLNWLDAQSGLASAPVSGAGNHXYCRNPDE 54
DB 127 CYEDGISTRGWSTAESGAECTNM--NSSALAQKPYSGRRPDALIRLGLGNHNYCRNPDR 184
QY 55 DPGFPCVYSGEAGVPEKRPCEDLRCPEPTTS 85
DB 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 213

RESULT 6
US-60-452-680-14405
; Sequence 14405, Application US/60452680
; GENERAL INFORMATION:


```

APPLICANT: CARGILL, Michele
APPLICANT: GRUPE, Andrew
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001450
CURRENT APPLICATION NUMBER: US/60/452,680
CURRENT FILING DATE: 2003-03-07
NUMBER OF SEQ ID NOS: 116213
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14405
LENGTH: 516
TYPE: PRT
ORGANISM: Homo sapiens
US-60-452-680-14405

Query Match      12.0%; Score 154.5; DB 7; Length 516;
Best Local Similarity 39.6%; Pred. No. 5.5e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

OY      4 CFWDNGLYREDQTSPPAPLRCLNMLDAOSGLASAPVS-----GAGNHSYCRNPDE 54
          :::::
DB      81 CYEDQGISIRGTWSTHSEASAECTNW--NSSALAQRKYSGRPRDAIRLGLGNHNYCRNPDR 138
          :::::

OY      55 DPGPWCYVSGEAGVPEKRPCEDLRCPETTS 85
          :::::
DB      139 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 167
          :::::

RESULT 7
US-60-453-135-8957
Sequence 8957, Application US/60453135
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
APPLICANT: IAKOUBOVA, Olga
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001456
CURRENT APPLICATION NUMBER: US/60/453,135
CURRENT FILING DATE: 2003-03-10
NUMBER OF SEQ ID NOS: 82762
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8957
LENGTH: 516
TYPE: PRT
ORGANISM: Homo sapiens
US-60-453-135-8957

Query Match      12.0%; Score 154.5; DB 7; Length 516;
Best Local Similarity 39.6%; Pred. No. 5.5e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

OY      4 CFWDNGLYREDQTSPPAPLRCLNMLDAOSGLASAPVS-----GAGNHSYCRNPDE 54
          :::::
DB      81 CYEDQGISIRGTWSTHSEASAECTNW--NSSALAQRKYSGRPRDAIRLGLGNHNYCRNPDR 138
          :::::

OY      55 DPGPWCYVSGEAGVPEKRPCEDLRCPETTS 85
          :::::
DB      139 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 167
          :::::

RESULT 8
US-60-453-050-8957
Sequence 8957, Application US/60453050
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
APPLICANT: LUKE, May
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001457
CURRENT APPLICATION NUMBER: US/60/453,050
CURRENT FILING DATE: 2003-03-10
NUMBER OF SEQ ID NOS: 82762
SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 8957
; LENGTH: 516
; TYPE: PR
; ORGANISM: Homo sapiens
US-60-453-050-8957

Query Match      12.0%; Score 154.5; DB 7; Length 516;
Best Local Similarity 39.6%; Pred.No. 5.5e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

OY      4 CFWDNGHL YREDQTS PAAGLRCLNMLDQSGLASPVS-----GAGNHSYCRNPDE 54
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      81 CYEDGISYRGTWSVAESGABECTNN--NSSALAQKPYSGRRPDALRLGLGNHNICRNPDR 138
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY      55 DPGPMCVYSGEAGVPEKRRPCEDLRCPEPTTS 85
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB     139 DSK-PMCTYF-KAGKYSSEFCSTPACSGNS 167

```

```

RESULT 9
US-09-612-314A-51
; Sequence 51, Application US/09612314A
; GENERAL INFORMATION:
; APPLICANT: SMITH, RICHARD ANTHONY GODWIN
; APPLICANT: DODD, IAN
; APPLICANT: MOSSAKOWSKA, DANUTA EWA IRENA
; TITLE OF INVENTION: CONJUGATES OF SOLUBLE PEPTIDIC COMPOUNDS WITH
; FILE OF INVENTION: MEMBRANE-BINDING AGENTS
; FILE REFERENCE: 37945-0004
; CURRENT APPLICATION NUMBER: US/09/612.314A
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 09/214,913
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: PCT/EP97/03715
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: GB 96 148 71.3
; PRIOR FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 51
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: tissue plasminogen activator
US-09-612-314A-51

Query Match          12.0%; Score 154.5; DB 5; Length 527;
Best Local Similarity 39.6%; Pred. No. 5,6e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4

QY      4 CFWNDGHLVREDPTSPAPGIRCLNMLDAQGLASAPVS-----GAGNHSYCRNPDE 54
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      92 CYVDGQISIRGTMSTABSGAECTNW--NSSALAQRKYSRRRDPATRLGNNHNYCRNPDR 149

QY      55 DPRGWCYVSGEAGVPEKAPCEDLRCPETTS 85
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      150 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 178

RESULT 10
PCT-US02-32263-26
; Sequence 26, Application PC/TUS0232263
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: DeFrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Bower, Caryn
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; TITLE OF INVENTION: REMODELING AND GLYCOCONJUGATION OF PEPTIDES
; FILE REFERENCE: 040853-01-5050WO

```

```

? CURRENT APPLICATION NUMBER: US02-10-31
? CURRENT FILING DATE: 2002-10-31
? PRIOR APPLICATION NUMBER: US 60/328,523
? PRIOR FILING DATE: 2001-10-10
? PRIOR APPLICATION NUMBER: US 60/344,692
? PRIOR FILING DATE: 2001-10-19
? PRIOR APPLICATION NUMBER: US 60/334,233
? PRIOR FILING DATE: 2001-11-08
? PRIOR APPLICATION NUMBER: US 60/334,301
? PRIOR FILING DATE: 2001-11-08
? PRIOR APPLICATION NUMBER: US 60/387,292
? PRIOR FILING DATE: 2002-06-07
? PRIOR APPLICATION NUMBER: US 60/391,777
? PRIOR FILING DATE: 2002-06-25
? PRIOR APPLICATION NUMBER: US 60/396,594
? PRIOR FILING DATE: 2002-07-17
? PRIOR APPLICATION NUMBER: US 60/404,249
? PRIOR FILING DATE: 2002-08-16
? PRIOR APPLICATION NUMBER: US 60/407,527
? PRIOR FILING DATE: 2002-08-28
? NUMBER OF SEQ ID NOS: 62
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 26
? LENGTH: 562
? TYPE: PRT
? ORGANISM: Homo sapiens
? PCT-US02-32263-26

```

Query Match	12.0%	Score 154.5;	DB 1;	Length 562;
Best Local Similarity	39.6%;	Pred. No. 6.1e-05;		
Matches	36;	Conservative	6;	Mismatches 36; Indels 13; Gaps 4;
QY	4	CPMNGHLVREDTSPARGLCNMLDAOSGLASPVS-----GAGNHSYCNPDE	54	
	:	:::	::	
Dd	127	CYEDQGISTYGTVSTASGAECTNM--NSSALAQKPYSGRRPDAIRLIGLNHYNCRPDR	184	
QY	55	DPRGPWCYVGSEAGVEKRPCEDLNCPETTS	85	
	:	: :		
Dd	185	DSK-PWCYVF-KAGKYSSEFCSTPAACSEGN	213	

```

RESULT 11
US-09-703-695A-4
; Sequence 4, Application US/09703695A
; GENERAL INFORMATION:
; APPLICANT: Xu, Yuan
; TITLE OF INVENTION: REVERSE-PHASE HPLC ASSAY FOR PLASMINOGEN ACTIVATORS
; FILE REFERENCE: P1788R1
; CURRENT APPLICATION NUMBER: US/09/703, 695A
; CURRENT FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: US 60/163,607
; PRIOR FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 4
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-703-695A-4

```

Query Match	Similarity	Score	DB	Length
Best Local	39.6%	Pred. No. 6.1e-05		562
Matches	36	Conservative	6	Mismatches 36, Indels 13, Gaps 4

QY	4	CFDNDHLTYREQOTSAPGLRCMLNWLDAQSLASAPVS-----GAGNHSYCNRPDE	54
	127	CFEDQGISIRGTWSTAESGAECTNN--NSSALACKPYSGRRPDALRIGLNHNHYCNPDR	184

QY	55	DPDGPWCYVSGEAGVPEKRRPCEDLRCPETTS	85
Db	185	DSK-PWCYVF-VAGYSESEFCTSPACSGNS	213

RESULT 12

```

US-10-287-994-26
; Sequence 26, Application US/10287994
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: Defrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Bower, Caryn
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; TITLE OF INVENTION: REMODELLING AND GLYCOCONJUGATION OF PEPTIDES
; FILE REFERENCE: 040853-01-5052-00
; CURRENT APPLICATION NUMBER: US/10/287,994
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-287-994-26

```

```
Query Match          12.0%; Score 154.5; DB 6; Length 562;
Best Local Similarity 39.6%; Pred. No. 6,1e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
```

Oy 4 CFWGDGHIYREDQTSPPAGLCRLMNLADQSGIASAPVS-----GAGNHSYCRNPDE 54

 |:::| |||||

Db 127 CYEDDGISIRGTWSTAESGABACTNW--NSSALAQKPYGGRRPDAIRLIGLNHNVCRNPD 184

Oy : |||||

Oy 55 DPRGMCVCSGEAGVPEKRKPCEDLARCPETTS 85

 |:::| |||||

Db 185 DSK--PWCYVF--KAGKYSEFCSTPACSGNS 213

```

RESULT 13
US-60-452-680-14408
; Sequence 14408, Application US/60452680
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01450
; CURRENT APPLICATION NUMBER: US/60/452,680
; CURRENT FILING DATE: 2003-03-07
; NUMBER OF SEQ. ID NOS: 116213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14408
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-60-452-680-14408

```

Query Match	12.0%	Score 154.5	DB 7	Length 562
Best Local Similarity	39.6%	Pred. No. 6.1e-05		
Matches 36; Conservative	6	Mismatches 36	Indels 13	Gaps 4
OY	4	CFWDNGHLYREDQTSPPARJRCINLWIDAOGLSAPVS-----GAGNHSYCRNPDE	54	

QY 4 CFWDNGHLYREDQTSPPGLRCLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 54

Db 127 CYEDGISTYRGWSTAESGAECTNW--NSSALAQKPYSGRRPDALRLGLGNHNYCRNPDR 184

Qy 55 DRGPMWCYVSGEAGVPEKRPCEDLRCPEPTTS 85
Db 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 213

RESULT 14

US-60-453-135-8960
; Sequence 8960, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8960
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-8960

Query Match 12.0%; Score 154.5; DB 7; Length 562;
Best Local Similarity 39.6%; Pred. No. 6.1e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

Qy 4 CFWDNGHLVREDQTSPPAPGLRCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 54
Db 127 CYEDGISTYRGWSTAESGAECTNW--NSSALAQKPYSGRRPDALRLGLGNHNYCRNPDR 184

Qy 55 DRGPMWCYVSGEAGVPEKRPCEDLRCPEPTTS 85
Db 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 213

RESULT 15

US-60-453-050-8960
; Sequence 8960, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8960
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-8960

Query Match 12.0%; Score 154.5; DB 7; Length 562;
Best Local Similarity 39.6%; Pred. No. 6.1e-05;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

Qy 4 CFWDNGHLVREDQTSPPAPGLRCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 54
Db 127 CYEDGISTYRGWSTAESGAECTNW--NSSALAQKPYSGRRPDALRLGLGNHNYCRNPDR 184

Qy 55 DRGPMWCYVSGEAGVPEKRPCEDLRCPEPTTS 85
Db 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 213

Search completed: April 7, 2003, 09:30:47

Job time : 37.2245 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:28 ; Search time 24.1497 Seconds
(without alignments)
294.842 Million cell updates/sec

Title: US-10-057-951-2_COPY_22_263
Perfect score: 1289
Sequence: 1 SGGCFWMDGHLVREDQTSPPA.....PVDPOEGSTPLMGAGTPGA 242

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_Aa: *
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/PTCUS.COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/backfilltest.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1289	100.0	263	4 US-09-411-977-2	Sequence 2, Appli
2	158.5	12.3	472	2 US-08-811-949-63	Sequence 63, Appl
3	154.5	12.0	437	2 US-08-811-949-49	Sequence 49, Appl
4	154.5	12.0	437	2 US-08-811-949-51	Sequence 51, Appl
5	154.5	12.0	437	2 US-08-811-949-55	Sequence 55, Appl
6	154.5	12.0	437	2 US-08-811-949-57	Sequence 57, Appl
7	154.5	12.0	527	1 US-07-609-510B-16	Sequence 16, Appl
8	154.5	12.0	527	1 US-08-811-949-39	Sequence 39, Appl
9	154.5	12.0	527	5 PCR-US91-01025A-2	Sequence 2, Appli
10	154.5	12.0	527	5 5185259-8	Sequence 39, Appl
11	154.5	12.0	527	6 5520913-1	Sequence 2, Appli
12	154.5	12.0	527	6 5200340-6	Sequence 2, Appli
13	154.5	12.0	562	2 US-08-811-949-43	Sequence 43, Appl
14	154.5	12.0	562	2 US-08-560-098A-50	Sequence 50, Appl
15	154.5	12.0	562	2 US-08-883-795A-38	Sequence 38, Appl
16	154.5	12.0	562	6 5185259-3	Sequence 39, Appl
17	154.5	12.0	562	6 5200340-2	Sequence 2, Appli
18	154.5	12.0	562	6 5344773-2	Sequence 2, Appli
19	148.5	11.5	83	2 US-08-811-949-2	Sequence 2, Appli
20	146.5	11.4	655	2 US-08-148-910-12	Sequence 12, Appl
21	146.5	11.4	655	1 US-08-448-937A-12	Sequence 12, Appl
22	142	11.0	160	2 US-08-612-788-35	Sequence 35, Appl
23	142	11.0	160	2 US-09-066-028-35	Sequence 35, Appl
24	142	11.0	250	2 US-08-612-788-30	Sequence 30, Appl
25	142	11.0	250	3 US-09-066-028-30	Sequence 30, Appl
26	142	11.0	339	1 US-08-248-629A-3	Sequence 3, Appli
27	142	11.0	339	1 US-08-451-932-3	Sequence 3, Appli

28	142	11.0	339	1 US-08-452-260-3	Sequence 3, Appli
29	142	11.0	339	1 US-08-326-785-3	Sequence 3, Appli
30	142	11.0	339	2 US-08-612-788-3	Sequence 3, Appli
31	142	11.0	339	2 US-08-605-598B-3	Sequence 3, Appli
32	142	11.0	339	2 US-08-429-743-3	Sequence 3, Appli
33	142	11.0	339	2 US-08-866-735-3	Sequence 3, Appli
34	142	11.0	339	3 US-09-066-028-3	Sequence 3, Appli
35	142	11.0	339	5 PCR-US95-05107-3	Sequence 3, Appli
36	142	11.0	352	2 US-08-612-788-40	Sequence 40, Appl
37	142	11.0	352	3 US-09-066-028-40	Sequence 40, Appl
38	142	11.0	374	4 US-09-377-250-3	Sequence 3, Appli
39	142	11.0	375	4 US-09-377-250-2	Sequence 2, Appli
40	142	11.0	378	2 US-08-612-788-42	Sequence 42, Appl
41	142	11.0	378	3 US-09-066-028-42	Sequence 42, Appl
42	142	11.0	378	4 US-09-206-059-1	Sequence 1, Appli
43	142	11.0	451	4 US-09-377-250-1	Sequence 1, Appli
44	142	11.0	452	4 US-09-377-250-4	Sequence 4, Appli
45	142	11.0	790	1 US-08-469-486-54	Sequence 54, Appli

ALIGNMENTS

RESULT 1
US-09-411-977-2
; Sequence 2, Application US/09411977
; Patent No. 6372473
; GENERAL INFORMATION:
; APPLICANT: Moore, Paul A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
; FILE REFERENCE: PF378P1
; CURRENT APPLICATION NUMBER: US/09/411, 977
; CURRENT FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: 09/084,491
; EARLIER FILING DATE: 1998-05-27
; EARLIER APPLICATION NUMBER: 60/048,000
; EARLIER FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-411-977-2

Query Match 100.0%; Score 1289; DB 4; Length 263;
Best Local Similarity 100.0%; Pred. No. 9.9e-129;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGGCFWMDGHLVREDQTSPPA...PVDPOEGSTPLMGAGTPGA 60
DB 22 SGGCFWMDGHLVREDQTSPPA...PVDPOEGSTPLMGAGTPGA 81
QY 61 CVYSGEAGVEKPCEDLRCPETTSQALPAFTTEIOEASGEGPADEVQVAPANALPARS 120
DB 82 CVYSGEAGVEKPCEDLRCPETTSQALPAFTTEIOEASGEGPADEVQVAPANALPARS 141
QY 121 EAAAVQPVISORVWMSKEKDLGLGVLTITMVIITAGIITIGSYKRGKDK 180
DB 142 EAAAVQPVISORVWMSKEKDLGLGVLTITMVIITAGIITIGSYKRGKDK 201
QY 181 EOHDOKVCEREMRITLPLSAFTNPCEIYDEKTVVHVSQTPVDPOEGSTPLMGAGTPGA 240
DB 202 EOHDOKVCEREMRITLPLSAFTNPCEIYDEKTVVHVSQTPVDPOEGSTPLMGAGTPGA 261
QY 241 GA 242
DB 262 GA 263

RESULT 2

```

US-08-811-949-63
; Sequence 63, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 472 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-811-949-63

Query Match 12.3%; Score 158.5; DB 2; Length 472;
Best Local Similarity 40.7%; Pred. No. 2,1e-08;
Matches 37; Conservative 5; Mismatches 36; Indels 13; Gaps 4

Cy 4 CFWMDGHLTREQOTSPAPGLRCINWIDAOGLASAPVS-----GAGNSYCNRPDE 54
Db 37 CYEDGISIRGWTMTAESAELECTNW--NSALAQRKYSRRPDRRLGLGINNYCNRPDR 94
Cy 55 DPRGWCYVSGEAGVPEKRPCEDDLRCPEFTTS 85
Db 95 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNIS 123

RESULT 3
US-08-811-949-49
; Sequence 49, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 472 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-811-949-63

```

```

1 STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
2 CITY: ARLINGTON
3 STATE: VA
4 COUNTRY: USA
5 ZIP: 22202
6
7 COMPUTER READABLE FORM:
8 MEDIUM TYPE: Floppy disk
9 COMPUTER: IBM PC compatible
10 OPERATING SYSTEM: PC-DOS/MS-DOS
11 SOFTWARE: Patentin Release #1.0, Version #1.30
12
13 CURRENT APPLICATION DATA:
14 APPLICATION NUMBER: US/08/811,949
15 FILING DATE: 05-MAR-1997
16 CLASSIFICATION: 435
17 ATTORNEY/AGENT INFORMATION:
18 NAME: ORLON, NORMAN F.
19 REGISTRATION NUMBER: 24,618
20 REFERENCE/DOCKET NUMBER: 18-966-0
21 TELECOMMUNICATION INFORMATION:
22 TELEPHONE: 703-413-3000
23 TELEFAX: 703-413-2220
24
25 INFORMATION FOR SEQ ID NO: 49:
26
27 SEQUENCE CHARACTERISTICS:
28 LENGTH: 437 amino acids
29 TYPE: amino acid
30 TOPOLOGY: linear
31 MOLECULE TYPE: protein
32
33 US-08-811-949-49
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
98
```


Patent No. 5326700
GENERAL INFORMATION:
APPLICANT: Bect et al.
TITLE OF INVENTION: Method for Altering Post-Translational Processing of Tissue Pl
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: IN.
COUNTRY: U.S.A.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/609,510B
FILING DATE: 19901106
CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-609-510B-16

Query Match 12.0%; Score 154.5; DB 1; Length 527;
Best Local Similarity 39.6%; Pred. No. 6,7e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 4 CFWDNGHLYREDQTSPPARCLNMLDQSGLASAPVS-----GAGNHSYCRNPDE 54
DB 92 CYEDQGISYRGWTSTAESGAECTNW--NSGALAQKYSGRPRDPIRLGLGNHNYCNPDR 149
QY 55 DPGPMCVYSGAGVPEKRPCELDRCPEPTS 85
DB 150 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 178

RESULT 8
US-08-811-949-39
Sequence 39, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIWA, MINO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
STREET: 1765 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-39

Query Match 12.0%; Score 154.5; DB 2; Length 527;
Best Local Similarity 39.6%; Pred. No. 6,7e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 4 CFWDNGHLYREDQTSPPARCLNMLDQSGLASAPVS-----GAGNHSYCRNPDE 54
DB 92 CYEDQGISYRGWTSTAESGAECTNW--NSGALAQKYSGRPRDPIRLGLGNHNYCNPDR 149
QY 55 DPGPMCVYSGAGVPEKRPCELDRCPEPTS 85
DB 150 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 178

RESULT 9
PCT-US91-01025A-2
Sequence 2, Application PCT/TUS9101025A
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
TITLE OF INVENTION: Tissue Plasminogen Activator Having Fibrin
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/01025A
FILING DATE: 19910214
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/486,657
FILING DATE: 1 March 1990
ATTORNEY/AGENT INFORMATION:
NAME: Haasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 454P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/266-1896
TELEFAX: 415/952-9881
TELEX: 910371-7168
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
PCT-US91-01025A-2

Query Match 12.0%; Score 154.5; DB 5; Length 527;
Best Local Similarity 39.6%; Pred. No. 6,7e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 4 CFMDNGHLTYREDQTSAPAGRLCLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 54
Db 92 CYEDQGISYRGTWSTAESGAECTNW--NSSALAQKPYSGRRPDAIRLIGLNHNHCNRPDR 149
QY 55 DPGPMCVYSGEAGVPEKRPCEDLRCPEETS 85
Db 150 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 178

RESULT 10

5185259-8
Patent No. 5185259
APPLICANT: GOEDEDEL, DAVID V.; KOHR, WILLIAM U.; PENNICA, DIANE;
VERAR, GORDON A.
TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN
ACTIVATOR
NUMBER OF SEQUENCES: 15
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/489,855
FILING DATE: 02-MAR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 12,694
FILING DATE: 09-FEB-1987
APPLICATION NUMBER: 483,052
FILING DATE: 07-APR-1983
APPLICATION NUMBER: 398,003
FILING DATE: 14-JUL-1982
APPLICATION NUMBER: 374,860
FILING DATE: 05-MAY-1982
SEQ ID NO: 8
LENGTH: 527

Query Match 12.0%; Score 154.5; DB 6; Length 527;
Best Local Similarity 39.6%; Pred. No. 6, 7e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 4 CFMDNGHLTYREDQTSAPAGRLCLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 54
Db 92 CYEDQGISYRGTWSTAESGAECTNW--NSSALAQKPYSGRRPDAIRLIGLNHNHCNRPDR 149
QY 55 DPGPMCVYSGEAGVPEKRPCEDLRCPEETS 85
Db 150 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 178

RESULT 11

5520913-1
Patent No. 5520913
APPLICANT: ANDERSON, STEPHEN; BENNETT, WILLIAM F.; ROTSTEIN,
DAVID; HIGGINS, DEBORAH L.; POKONI, NICHOLAS F.; ZOLLER, MARK J.
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR HAVING
ZYMOGENIC PROPERTIES
NUMBER OF SEQUENCES: 35
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/88,451
FILING DATE: 06-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 770,510
FILING DATE: 03-OCT-1991
APPLICATION NUMBER: 384,608
FILING DATE: 24-JUL-1989
APPLICATION NUMBER: 240,856
FILING DATE: 02-SEP-1988
SEQ ID NO: 1
LENGTH: 527

Query Match 12.0%; Score 154.5; DB 6; Length 527;
Best Local Similarity 39.6%; Pred. No. 6, 7e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 4 CFMDNGHLTYREDQTSAPAGRLCLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 54
Db 92 CYEDQGISYRGTWSTAESGAECTNW--NSSALAQKPYSGRRPDAIRLIGLNHNHCNRPDR 149
QY 55 DPGPMCVYSGEAGVPEKRPCEDLRCPEETS 85
Db 150 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 178

RESULT 12

5200340-6
Patent No. 5200340
APPLICANT: FOSTER, DONALD C.; MULVIHILL, EILEEN R.; O'HARA,
PATRICK J.; PINGEL, KURT; YOSHITAKE, SHINJI
TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
ACTIVATORS
NUMBER OF SEQUENCES: 34
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/53,412
FILING DATE: 22-MAY-1987
SEQ ID NO: 6
LENGTH: 546

Query Match 12.0%; Score 154.5; DB 6; Length 546;
Best Local Similarity 39.6%; Pred. No. 7e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 4 CFMDNGHLTYREDQTSAPAGRLCLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 54
Db 127 CYEDQGISYRGTWSTAESGAECTNW--NSSALAQKPYSGRRPDAIRLIGLNHNHCNRPDR 184
QY 55 DPGPMCVYSGEAGVPEKRPCEDLRCPEETS 85
Db 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGNS 213

RESULT 13

US-08-811-949-43
Sequence 43, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIMA, MINO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUTI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESS: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000

TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 562 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-43

Query Match 12.0%; Score 154.5; DB 2; Length 562;
Best Local Similarity 39.6%; Pred. No. 7.3e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 4 CFWNGHLYREDQTSAPAGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 54
DB 127 CYEDGISTYRGWTSTAGSAGACTNM--NSSALAKPYSGRRPDAIRLGIGNHNYCRNPDR 184
QY 55 DPRGWCYSGEAGVPEKRPCEDLACPETTS 85
DB 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 213

RESULT 14
US-08-560-098A-50
Sequence 50, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WENDET, Stephan
APPLICANT: HEINZEL, WILFRIED, Regina
TITLE OF INVENTION: Proteins having Fibrinolytic and
NUMBER OF INVENTION: Coagulation-Inhibiting Properties
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 562 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-560-098A-50

Query Match 12.0%; Score 154.5; DB 2; Length 562;
Best Local Similarity 39.6%; Pred. No. 7.3e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;
QY 4 CFWNGHLYREDQTSAPAGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 54
DB 127 CYEDGISTYRGWTSTAGSAGACTNM--NSSALAKPYSGRRPDAIRLGIGNHNYCRNPDR 184

DB 127 CYEDGISTYRGWTSTAGSAGACTNM--NSSALAKPYSGRRPDAIRLGIGNHNYCRNPDR 184
QY 55 DPRGWCYSGEAGVPEKRPCEDLACPETTS 85
DB 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 213

RESULT 15
US-08-883-795A-38
Sequence 38, Application US/08883795A
Patent No. 5985607
GENERAL INFORMATION:
APPLICANT: Delcave, Genevieve
APPLICANT: Awang, Gregor
TITLE OF INVENTION: Recombinant DNA Molecules and Expression
NUMBER OF INVENTION: Vectors for Tissue Plasminogen Activator
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: BRESKIN & PARR
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/883,795A
FILING DATE: 27-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gravelle, Micheline
REGISTRATION NUMBER: 40,261
REFERENCE/DOCKET NUMBER: 7841-062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 562 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein (tpa)
US-08-883-795A-38

Query Match 12.0%; Score 154.5; DB 2; Length 562;
Best Local Similarity 39.6%; Pred. No. 7.3e-08;
Matches 36; Conservative 6; Mismatches 36; Indels 13; Gaps 4;

QY 4 CFWNGHLYREDQTSAPAGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 54
DB 127 CYEDGISTYRGWTSTAGSAGACTNM--NSSALAKPYSGRRPDAIRLGIGNHNYCRNPDR 184
QY 55 DPRGWCYSGEAGVPEKRPCEDLACPETTS 85
DB 185 DSK-PWCYVF-KAGKYSSEFCSTPACSEGN 213

Search completed: April 7, 2003, 09:22:29
Job time : 26.1497 sec

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:25 ; Search time 4.86486 Seconds

(without alignments)
1185.658 Million cell updates/sec

Title: US-10-057-951-2_COPY_25_84

Perfect score: 357
Sequence: 1 CFMDNGHLVREDQTSAPAGL.....GNHSYCRNPDEDPGRWCYV 60Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_73.*

1: p1r1.*
2: p1r2.*
3: p1r3.*
4: p1r4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	146	40.9	562 1 UKHUT	t-plasminogen acti
2	141	39.5	559 1 A35029	t-plasminogen acti
3	141	39.5	655 1 A46688	hepatocyte growth
4	137	38.4	291 2 I38098	t-plasminogen acti
5	136	38.1	559 1 A29941	t-plasminogen acti
6	134	37.5	431 2 JS0599	t-plasminogen acti
7	134	37.5	477 1 A34369	t-plasminogen acti
8	134	37.5	477 2 JS0598	t-plasminogen acti
9	132	37.0	433 1 UKBAY	u-plasminogen acti
10	131	36.7	431 1 UKHU	u-plasminogen acti
11	130	36.4	442 1 UKPG	u-plasminogen acti
12	128	35.9	716 1 JCS061	u-plasminogen acti
13	127	35.6	433 1 JN0560	macrophage-stimula
14	126	35.3	394 2 JS0600	u-plasminogen acti
15	126	35.3	716 1 A40332	t-plasminogen acti
16	123	34.5	169 2 A40522	macrophage-stimula
17	121	33.9	810 2 PLHU	plasmin (EC 3.4.21
18	119	33.3	433 1 UKMS	u-plasminogen acti
19	119	33.3	728 1 JH0579	hepatocyte growth
20	118	33.1	432 1 SI8932	u-plasminogen acti
21	118	33.1	728 1 A60185	hepatocyte growth
22	118	33.1	810 2 I46260	plasmin (EC 3.4.21
23	115	32.2	477 2 JS0597	t-plasminogen acti
24	113.5	31.8	790 1 PLPG	plasmin (EC 3.4.21
25	113.5	31.8	4548 1 JS0657	apoptoteicn(a) (EC
26	113	31.7	810 2 B30848	plasmin (EC 3.4.21
27	112	31.4	728 1 A35644	hepatocyte growth
28	111.5	31.2	603 2 S28941	coagulation factor
29	111	31.1	812 1 PLBO	plasmin (EC 3.4.21

30	109	30.5	685 1 A40289	neurotrophic recep
31	109	30.5	711 1 A47136	macrophage-stimula
32	108	30.3	625 1 TBBO	thrombin (EC 3.4.2
33	107.5	30.1	123 2 B61545	plasmin (EC 3.4.21
34	106.5	29.8	434 1 A35005	u-plasminogen acti
35	105.5	29.6	460 2 B61545	plasmin (EC 3.4.21
36	105	29.4	622 1 TBHU	thrombin (EC 3.4.2
37	105	29.4	710 1 I51283	hepatocyte growth
38	104.5	29.3	812 1 PLMS	plasmin (EC 3.4.21
39	104	29.1	560 1 JC4795	plasma hyaluronan-
40	101.5	28.4	120 2 B61545	plasmin (EC 3.4.21
41	100.5	28.2	455 2 A61545	plasmin (EC 3.4.21
42	98.5	27.6	593 2 S45281	coagulation factor
43	97	27.2	617 2 SI0511	thrombin (EC 3.4.2
44	95.5	26.8	89 2 A60140	plasmin (EC 3.4.21
45	95	26.6	618 2 A35827	thrombin (EC 3.4.2

ALIGNMENTS

RESULT 1

UKHUT

t-plasminogen activator (EC 3.4.21.68) precursor [validated] - human
N:Alternate names: t-PA; tissue plasminogen activator
C/Species: Homo sapiens (man)

C/Date: 14-Nov-1983 #sequence revision 14-Nov-1983 #text change 08-Dec-2000

C/Accession: A94004; A23529; J70562; A93293; S02125; A91343; A93951; A91322; A54645; I60

R/NY, T.; Eligh, F.; Lund, B.
Proc. Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984A:/Title: The structure of the human tissue-type plasminogen activator gene: correlation
A:/Reference number: A94004; MUID:84298137; PMID:6089198

A:/Accession: A94004

A:/Molecule type: DNA

A:/Residues: 1-562 <NVT>

A:/Cross-references: GB:I00141

A:/Note: the codon given for residue 93 (ACC) is inconsistent with the authors' translati.

R:/Friedman Degen, S.J.; Rajput, B.; Retch, E.

J. Biol. Chem. 261, 6972-6985, 1986

A:/Title: The human tissue plasminogen activator gene.

A:/Reference number: A23529; MUID:86196143; PMID:3009482

A:/Accession: A23529

A:/Molecule type: DNA

A:/Residues: 1-562 <DEG>

A:/Cross-references: GB:K03021; NID:G339817; PIDN:AAA98809.1; PID:G339818

R:/Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuda, S.; Higashio, K.

Agric. Biol. Chem. 55, 1225-1232, 1991

A:/Title: Purification and characterization of tissue plasminogen activator secreted by h

A:/Reference number: J70562; MUID:91291340; PMID:1368681

A:/Accession: J70562

A:/Molecule type: mRNA

A:/Residues: 31-562 <ITR>

A:/Cross-references: DDBJ:D01096; NID:G220128; PIDN:BA00881.1; PID:G441174

A:/Experimental source: embryonic lung fibroblast IMR-90 cells

A:/Note: Part of this sequence, including the amino end of the mature protein, was confir

R:/Pentica, D.; Holmes, W.E.; Kohr, W.T.; Hartins, R.N.; Vehar, G.A.; Ward, C.A.; Bennett

Nature 301, 214-221, 1983

A:/Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Esche

A:/Reference number: A93293; MUID:83115262; PMID:6337343

A:/Accession: A93293

A:/Molecule type: mRNA

A:/Residues: 1-562 <PEN>

A:/Cross-references: GB:I00141

A:/Experimental source: melanoma cells

R:/Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M.

Nucleic Acids Res. 16, 5695, 1988

A:/Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human fe

A:/Reference number: S02125; MUID:88262579; PMID:3133640

A:/Accession: S02125

A:/Status: translation not shown

A:/Molecule type: mRNA

A:/Residues: 1-562 <SAS>

A:/Cross-references: EMBL:X07393; NID:G37243; PIDN:CMA0302.1; PID:G37244

A:Experimental source: fetal lung cells
 R:Kakitani, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Okada, K.; Ma
 FBBS Lett. 189, 145-149, 1988
 A:Title: Expression in Escherichia coli of finger-domain lacking tissue-type plasminogen
 A:Reference number: A91343; PMID:85285620; PMID:3896853
 A:Accession: A91343
 A:Molecule type: mRNA
 A:Residues: 1-38, 'G', 86-433, 'E', 435-562 <KRG>
 A:Experimental source: Detroit 562 cells; ATCC 138
 R:Edlund, T.; Ny, T.; Randy, M.; Heden, L.O.; Palm, G.; Holmgren, E.; Josephson, S.
 Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983
 A:Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen activator
 A:Reference number: A93951; PMID:83169656; PMID:6572897
 A:Accession: A93951
 A:Molecule type: mRNA
 A:Residues: 251-358 <EDL>
 A:Experimental source: melanoma cells
 R:Pohl, G.; Kallstrom, M.; Bergsdorf, N.; Wallen, P.; Jornvall, H.
 Biochemistry 23, 3701-3707, 1984
 A:Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived amino
 differences.
 A:Reference number: A90488; PMID:85000468; PMID:6433976
 A:Contents: annotation; melanoma cells, partial sequence of residues 36-562, active and
 R:Pohl, G.; Kaplan, L.; Biharsson, M.; Wallen, P.; Jornvall, H.
 FBBS Lett. 168, 29-32, 1984
 A:Title: Differences between uterine and melanoma forms of tissue plasminogen activator.
 A:Reference number: A91322; PMID:84158956; PMID:6538514
 A:Accession: A91322
 A:Molecule type: protein
 A:Residues: 33-45;311-320 <POH>
 A:Experimental source: uterus
 A>Note: in the uterus, cleavage of the activation peptide may also occur after 38-Gln
 R:van Zonneveld, A.J.; Veerman, H.; Pannekoek, H.
 J. Biol. Chem. 261, 14214-14218, 1986
 A:Reference number: A37567; PMID:87033611; PMID:3021732
 A:Contents: annotation; fibrin binding site
 R:Verheijen, J.H.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, P.H.; Enger
 EMBO J. 5, 3525-3530, 1986
 A:Title: Involvement of finger domain and kringle 2 domain of tissue-type plasminogen ac
 A:Reference number: A37568; PMID:87161761; PMID:3030730
 A:Contents: annotation; fibrin binding site
 R:Dodd, I.; Nunn, B.; Robinson, J.H.
 Thromb. Haemost. 59, 523-528, 1988
 A:Title: Isolation, identification and pharmacokinetic properties of human tissue-type F
 A:Reference number: A60902; PMID:89044681; PMID:3142086
 A:Contents: annotation; novel forms of expressed recombinant t-PA
 R:Harris, T.J.R.; Patel, T.; Marston, F.A.O.; Little, S.; Emtage, J.S.; Opdenakker, G.;
 Mol. Biol. Med. 3, 279-292, 1986
 A:Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its expr
 A:Reference number: A54645; PMID:86284200; PMID:3090401
 A:Accession: A54645
 A:Molecule type: mRNA
 A:Residues: 1-562 <HAR>
 A:Cross-references: GB:M15518; NID:G190031; PIDN:AAA60111.1; PID:G190032
 A>Note: parts of this sequence were confirmed by peptide sequencing
 R:Reddy, V.B.; Garramone, A.J.; Sasek, H.; Wei, C.
 DNA 6, 461-472, 1987
 A:Title: Expression of human uterine tissue-type plasminogen activator in mouse cells us
 A:Reference number: I60110; PMID:88054470; PMID:2824147
 A:Accession: I60110
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-562 <RES>
 A:Cross-references: GB:M18182; NID:G340176; PIDN:AAA36800.1; PID:G340177
 R:Fisher, R.; Waller, E.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schleuning, W.D.
 J. Biol. Chem. 260, 11223-11230, 1985
 A:Title: Isolation and characterization of the human tissue-type plasminogen activator s
 A:Reference number: I55232; PMID:85289338; PMID:3161893
 A:Accession: I55232
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-36 <RE2>
 A:Cross-references: GB:M11890; NID:G339837; PIDN:AAA61213.1; PID:G339839

C:Comment: Cleavage by plasmin or trypsin produces two chains held together by a single
 C:Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val bond. It
 C:Comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repeat.
 C:Genetics:
 A:Gene: GDB:PLAT
 A:Cross-references: GDB:119496; OMIM:173370
 A:Map position: 8p12-8p12
 A:Introns: 24/3; 39/1; 122/1; 180/2; 211/1; 268/2; 297/1; 362/2; 408/1; 455/3; 510/
 C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
 C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase
 F:1-23/Domain: signal sequence #status predicted <PRO>
 F:24-32/Domain: signal sequence #status predicted <PRO>
 F:33-562/Product: t-plasminogen activator #status experimental <MAT>
 F:33-310/Product: t-plasminogen activator chain A #status experimental <ACH>
 F:41-78/Domain: fibronectin type I repeat homology <1F1>
 F:86-119/Domain: EGF homology <EGF>
 F:127-208/Domain: kringle homology <KR1>
 F:215-296/Domain: kringle homology <KR2>
 F:311-562/Product: t-plasminogen activator chain B #status experimental <BCH>
 F:311-556/Domain: trypsin homology <TRY>
 F:41-71,68-78,86-97,91-108,110-119,127-208,148-190,179-203,215-226,236-278,267-291,299-4
 F:152,483/Binding site: carbohydrate (asn) (covalent) #status experimental
 F:219/Binding site: carbohydrate (asn) (covalent) (partial) #status experimental
 F:310-311/Cleavage site: Arg-116 (plasmin, trypsin) #status experimental
 F:357,406/Active site: His, Asp #status predicted
 F:513/Active site: Ser #status experimental
 Query Match 40.9%; Score 146; DB 1; Length 562;
 Best Local Similarity 43.5%; Pred. No. 7,4e-09;
 Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;
 QY 1 CPMDGHLVREDQTSAPGRCLNWLDAQSGLASAPVS-----GAGNYSYCNPDE 51
 DB 127 CYPDGISTRICTWSTASGAECTNW--NSSALAKPFSGRPRDARLGLGNHYCNRPDR 184
 QY 52 DPGPMCYV 60
 DB 185 DSK-PWCYV 192
 RESULT 2
 A35029
 C:Plasminogen activator (BC 3.4.21.68) precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A35029; A31597
 R:Feng, P.; Ohlsson, M.; Ny, T.
 J. Biol. Chem. 265, 2022-2027, 1990
 A:Title: The structure of the TARA-less rat tissue-type plasminogen activator gene. Spec
 A:Reference number: A35029; PMID:90130448; PMID:2105315
 A:Accession: A35029
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-559 <FEN>
 A:Cross-references: GB:M31197; NID:G207429; PIDN:AAA42261.1; PID:G207431; GB:J05226
 R:Ny, T.; Leonardsson, G.; Heuvel, A.J.W.
 DNA 7, 671-677, 1988
 A:Title: Cloning and characterization of a cDNA for rat tissue-type plasminogen activato
 A:Reference number: A31597; PMID:89170114; PMID:3148445
 A:Accession: A31597
 A:Molecule type: mRNA
 A:Residues: 1-379, 'K', 381-559 <NYT>
 A:Cross-references: GB:M23697; NID:G530159; PIDN:AAA41812.1; PID:G530160
 C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
 C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
 F:1-17/Domain: signal sequence #status predicted <PRO>
 F:18-29/Domain: propeptide #status predicted <PRO>
 F:30-559/Product: t-plasminogen activator #status predicted <MAT>
 F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
 F:38-75/Domain: fibronectin type I repeat homology <1F1>
 F:83-116/Domain: EGF homology <EGF>
 F:124-205/Domain: kringle homology <KR1>
 F:213-294/Domain: kringle homology <KR2>

F:309-559/Domain: t-Plasminogen activator chain B #status predicted <BCH>
F:309-553/Domain: t-Plasminogen activator chain B #status predicted <BCH>
F:36-68/66-75/83-94/88-105/107-116/124-205/145-187/176-200/213-294/234-276/265-289/297-4
F:149/481/Binding site: carboxylate (Asn) (covalent) #status predicted
F:308-309/Cleavage site: Arg-Ile (Plasmin, trypsin) #status predicted
F:355/404/510/Active site: His, Asp, Ser #status predicted

Query Match 39.5%; Score 141; DB 1; Length 559;
Best Local Similarity 42.0%; Pred. No. 2.7e-08;
Matches 29; Conservative 6; Mismatches 22; Indels 12; Gaps 3;

Qy 1 CFMDNGHLVREDQTPAPGRLCLNM-----LDAQSGLASPVSGAGNSYCRNPDE 51
Db 124 CFEGGITYRGWSTAGNACTNM--NSSALQKPYKAPRRPNAIKGLGNHNYCRNPDR 181

Qy 52 DPGPWCYV 60
Db 182 DYK-PWCYV 189

RESULT 3

A:Accession: A46688
A:Map position: 8p12-Bp12
A:Insertions: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: alternative splicing; fibrinolysis; glycoprotein; kringle
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-32/Domain: propeptide #status predicted <PRO>
F:33-251/Product: t-Plasminogen activator, inactive endothelial splice form #status pred
F:41-78/Domain: fibronectin type I repeat homology <1FA>
F:86-119/Domain: EGF homology <EGF>
F:127-208/Domain: kringle homology <KR1>
F:215-291/Domain: kringle homology #status atypical <KR2>
F:41-78/69-78/86-97/91-108/110-119/127-208/148-190/179-203/Diulfide bonds: #status pred

A:Residues: 1-655 <M1>
A:Cross-references: DDBJ:D14012; NID:G219680; PDB:BA03113.1; PID:G219681
A:Experimental source: Liver (mRNA); serum (protein)
A:Note: Sequence extracted from NCBI backbone (NCBI:11227, NCBI:P.131228)
A:Note: parts of the sequence, including the amino ends of the heavy and light chains, c
C:Genetics:
A:Gene: GDB:HGFAC; HGFA; HGFAP
A:Cross-references: GDB:9954514
A:Map position: 4p16-4p16
C:Function:
A:Description: activates hepatocyte growth factor by specific proteolytic cleavage
A:Pathway: tissue repair and regeneration
C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;
C:Keywords: glycoprotein; hydrolase; kringle; liver; plasma; serine proteinase
F:1-34/Domain: signal sequence #status predicted <SIG>
F:108-148/Domain: fibronectin type II repeat homology <1F2>
F:164-197/Domain: EGF homology <EG1>
F:202-237/Domain: fibronectin type I repeat homology <1F1>
F:245-278/Domain: EGF homology <EG2>
F:286-367/Domain: kringle homology <KR3>

F:373-407/Product: hepatocyte growth factor activator light chain #status experimental
F:408-655/Product: hepatocyte growth factor activator heavy chain #status experimental
F:408-641/Domain: trypsin homology <TRY>
F:408-48/290/468/492/546/Binding site: carboxylate (Asn) (covalent) #status predicted
F:164-175/169-186/188-197/202-230/228-237/245-256/250-267/269-278/286-367/307-349/338-36
F:447/497/599/Active site: His, Asp, Ser #status predicted

Query Match 39.5%; Score 141; DB 1; Length 655;
Best Local Similarity 45.6%; Pred. No. 3.2e-08;
Matches 31; Conservative 6; Mismatches 21; Indels 10; Gaps 3;

Qy 1 CFMDNGHLVREDQTPAPGRLCLNM-----LDAQSGLASPVSGAGNSYCRNPDE 52
Db 286 CFEGGITYRGWSTAGNACTNM--NSSALQKPYKAPRRPNAIKGLGNHNYCRNPDR 344

Qy 53 DPGPWCYV 60
Db 345 ER-PWCYV 351

RESULT 4

I38098
t-Plasminogen activator precursor, inactive endothelial splice form - human
N/Alternate names: tissue plasminogen activator
C:Species: Homo sapiens (man)
C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
C:Accession: I38098; S01678
R:Siebert, P.D.; Fong, K.

Nucleic Acids Res. 18, 1086, 1990
A:Title: Variant tissue type plasminogen activator (PLAT) cDNA obtained from human endoc
A:Reference number: I38098; MUID:90192128; PMID:1969145
A:Accession: I38098

A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-291 <S1E>
A:Cross-references: EMBL:X13097; NID:G35282; PDB:CAA31489.1; PID:G35283
C:Comment: For the main splice form, see FIR:OKHOT. This form probably does not have pro

C:Genetics:
A:Gene: GDB:PLAT
A:Cross-references: GDB:119496; OMIM:173370
A:Map position: 8p12-Bp12

A:Insertions: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: alternative splicing; fibrinolysis; glycoprotein; kringle
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-32/Domain: propeptide #status predicted <PRO>
F:33-251/Product: t-Plasminogen activator, inactive endothelial splice form #status pred
F:41-78/Domain: fibronectin type I repeat homology <1FA>
F:86-119/Domain: EGF homology <EGF>
F:127-208/Domain: kringle homology <KR1>
F:215-291/Domain: kringle homology #status atypical <KR2>
F:41-78/69-78/86-97/91-108/110-119/127-208/148-190/179-203/Diulfide bonds: #status pred

Query Match 38.4%; Score 137; DB 2; Length 291;
Best Local Similarity 42.0%; Pred. No. 4e-08;
Matches 29; Conservative 5; Mismatches 23; Indels 12; Gaps 3;

Qy 1 CFMDNGHLVREDQTPAPGRLCLNM-----LDAQSGLASPVSGAGNSYCRNPDE 51
Db 127 CFEGGITYRGWSTAGNACTNM--NSSALQKPYKAPRRPNAIKGLGNHNYCRNPDR 184

Qy 52 DPGPWCYV 60
Db 185 DSK-PWCYV 192

RESULT 5

A29941
t-Plasminogen activator (EC 3.4.21.68) precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A29941; S48205; S48207; S48206
R:Rickles, R.J.; Darrow, A.L.; Strickland, S.

J. Biol. Chem. 263, 1563-1569, 1988
A:Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mR
A:Reference number: A29941; MUID:88087303; PMID:2826484
A:Accession: A29941

A:Molecule type: mRNA
A:Residues: 1-559 <RC>
A:Cross-references: GB:J03520; NID:G202109; PDB:AAA0470.1; PID:G202110
R:Uhlen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.

Eur. J. Biochem. 224, 863-871, 1994
A:Title: Characterization of the murine plasma fibrinolytic system.
A:Reference number: S48202; MUID:95010076; PMID:7523120
A:Accession: S48205

A:Molecule type: protein
A:Residues: 33-37; 'X', 39-40 <LIU>
A:Accession: S48207

A:Molecule type: protein
A:Residues: 309-316 <LI2>
A:Accession: S48206
A:Molecule type: protein
A:Residues: 33-37; 'X', 39-40 <LIW>

A:Reference number: A51255; PDB:1KDU
 A:Contents: annotation, conformation and disulfide bond assignments by (1)H-NMR, residue
 RjLi, X.; Smith, R.A.G.; Dobson, C.M.
 Biochemistry 31, 9562-9571, 1992
 A>Title: Sequential (1)H NMR assignments and secondary structure of the kringle domain B
 A:Reference number: A44375; MUID:93003110; PMID:1327118
 A:Contents: annotation, conformation and disulfide bond assignments by (1)H-NMR
 R; Hansen, A.P.; Petros, A.M.; Meadows, R.P.; Nettesheim, D.G.; Mazur, A.P.; Olejniczak,
 submitted to the Brookhaven Protein Data Bank, January 1994
 A:Reference number: A66822; PDB:1HRK
 A:Contents: annotation, conformation and disulfide bond assignments by (1)H-NMR, residue
 R; Serragou, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobson, C.M.;
 submitted to the Brookhaven Protein Data Bank, July 1995
 A:Reference number: A66058; PDB:11MM
 A:Contents: annotation, X-ray crystallography, 2.5 angstroms, residues 168-175, 179-426
 C:Comment: This enzyme is found in urine in a high molecular mass form, consisting of A
 C:Comment: urokinase-type plasminogen activator proteolytically activates plasminogen, a
 C:Genetic: 8
 A:Gene: GDB:PLAU
 A:Cross-references: GDB:119497; OMIM:191840
 A:Map position: 10q24-10q24
 A:Insertions: 19/3; 23/1; 65/1; 123/2; 154/1; 227/2; 277/1; 324/2; 373/3
 C:Function:
 A:Description: proteolytically activates plasminogen
 A:Pathway: fibrinolysis
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C:Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine protease
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-431/Product: urokinase-type plasminogen activator, single chain form #status predic
 F:21-177/Product: urokinase-type plasminogen activator, chain A #status experimental <MR>
 F:31-62/Domain: EGF homology <EGF>
 F:70-151/Domain: kringle homology <KR>
 F:168-177/Product: urokinase-type plasminogen activator chain A1 #status experimental <M
 F:179-431/Product: urokinase-type plasminogen activator chain B #status experimental <M
 F:179-419/Domain: trypsin homology <TRY>
 F:31-33,33-51,53-62,70-151,91-133,132-146,168-299,209-225,217-288,313-382,345-361,372-40
 F:33/Binding site: carbohydrate (Thr) (covalent) #status predicted
 F:178-179/Cleavage site: Lys-Ile (plasmin) #status experimental
 F:224,275,376/Active site: His, Asp, Ser #status experimental
 F:332/Binding site: carbohydrate (Asn) (covalent) #status experimental
 Query Match 36.4%; Score 131; DB 1; Length 431;
 Best Local Similarity 40.3%; Pred. No. 2,9e-07;
 Matches 27; Conservative 7; Mismatches 25; Indels 8; Gaps 2;
 QY 1 CFPDNGHLYREDQTSFPAQLACLMWDA-----QSGSLASAPVSGAGNSYCNPDDEP 53
 Db 70 CFEQNGHFRKGASTDTWGRPLPMSATVLTQTHAHRSDALQGLGKHNVCNPD-NR 128
 QY 54 RGPWCYV 60
 Db 129 RRPWCYV 135
 RESULT 11
 UKPG
 u-plasminogen activator (EC 3.4.21.73) precursor - pig
 N:Alternate names: uPA
 C:Species: Sus scrofa domestica (domestic pig)
 C>Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998
 C:Accession: A00932
 R;Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.
 Nucleic Acids Res. 12, 9525-9541, 1984
 A>Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.
 A:Reference number: A00932; MUID:85087954; PMID:6096832
 A:Accession: A00932
 A:Molecule type: DNA
 A:Residues: 1-240, 'H', 242-442 <NA>
 A:Experimental source: kidney cell line LLC-PK1
 R;Nagamine, Y.
 submitted to the Protein Sequence Database, December 1986
 A:Reference number: A37566
 A:Contents: annotation; correction to residue 241

C:Genetic: 8
 A:Insertions: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine protease
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F:31-64/Domain: EGF homology <EGF>
 F:72-153/Domain: kringle homology <KR>
 F:190-442/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
 F:190-430/Domain: trypsin homology <TRY>
 F:152/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:179-310,220-236,228-299,324-393,356-372,383-411/Disulfide bonds: #status predicted
 F:235,286,387/Active site: His, Asp, Ser #status predicted
 Query Match 36.4%; Score 130; DB 1; Length 442;
 Best Local Similarity 44.8%; Pred. No. 3,9e-07;
 Matches 30; Conservative 4; Mismatches 25; Indels 8; Gaps 3;
 QY 1 CFPDNGHLYREDQTSFPAQLACLMWDAQSGL---ASAPVSGAGNSYCNPDDEP 53
 Db 72 CFEQNGHFRKGASTDTWGRPLPMSATVLTQTHAHRSDALQGLGKHNVCNPD-NQ 130
 QY 54 RGPWCYV 60
 Db 131 RRPWCYV 137
 RESULT 12
 JC5061
 macrophage-stimulating protein 1 precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 16-Jun-2000
 C:Accession: JC5061
 R;Ohnishi, K.; Iwama, A.; Matsuno, K.; Ezaki, T.; Sekamoto, O.; Hamaguchi, I.; Takasu, N
 Biochem. Biophys. Res. Commun. 227, 273-280, 1996
 A>Title: Molecular cloning of rat macrophage-stimulating protein and its involvement in
 A:Reference number: JC5061; MUID:97011126; PMID:8858136
 A:Accession: JC5061
 A:Molecule type: mRNA
 A:Residues: 1-716 <OHS>
 A:Cross-references: EMBL:X95096; NID:G1669718; PID:CA64473.1; PID:G1669719
 C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
 C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
 C:Keywords: duplication; glycoprotein; growth factor; kringle
 F:1-111/Domain: signal sequence #status predicted <SIG>
 F:312-488,489-716/Product: macrophage-stimulating protein 1 #status predicted <MAT>
 F:312-488/Domain: macrophage-stimulating protein 1 alpha chain #status predicted <ACH>
 F:110-186/Domain: kringle homology <KR1>
 F:191-268/Domain: kringle homology <KR12>
 F:292-370/Domain: kringle homology <KR13>
 F:379-457/Domain: kringle homology <KR14>
 F:489-716/Domain: macrophage-stimulating protein 1 beta chain #status predicted <BCH>
 F:489-709/Domain: trypsin homology <TRY>
 F:72,305,620/Binding site: carbohydrate (Asn) (covalent) #status predicted
 Query Match 35.9%; Score 128; DB 1; Length 716;
 Best Local Similarity 41.9%; Pred. No. 1,1e-06;
 Matches 26; Conservative 3; Mismatches 25; Indels 4; Gaps 2;
 QY 1 CFPDNGHLYREDQTSFPAQLACLMW---LDAQSLASAPVSGAGNSYCNPDDEP 57
 Db 110 CIMPNGHSTRGVARTKADGLPCQAMSRFPNDHKYTPTRKGL-ENFCRNPDDPGRPW 168
 QY 58 CY 59
 Db 169 CY 170
 RESULT 13
 JN0560
 u-plasminogen activator (EC 3.4.21.73) precursor - bovine
 N:Alternate names: uPA
 C:Species: Bos primigenius taurus (cattle)

C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
 C/Accession: JN0560
 R/Kraetschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schlenning, W.D.
 Gene 125, 177-183, 1993
 A>Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and induc
 A/Reference number: JN0560; PMID:93216119; PMID:8385052
 A/Accession: JN0560
 A/Molecule type: mRNA
 A/Residues: 1-433 <KRA>
 A/Cross-references: GB:U03546; NID:G163800; PIDN:AAA51419.1; PID:G163801
 C/Superfamily: urokinase-type plasminogen activator; BGF homology; kringie homology; try
 C/Keywords: glycoprotein; heterodimer; hydrolase; kringie; serine proteinase
 F:1-10/Domain: signal sequence #status predicted <SIG>
 F:21-119/Product: plasminogen activator chain A #status predicted <MA1>
 F:21-119/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F:33-64/Domain: BGF homology <EGF>
 F:72-153/Domain: kringie homology <KR3>
 F:181-433/Product: plasminogen activator chain B #status predicted <MA2>
 F:181-421/Domain: trypsin homology <TRY>
 F:170-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted
 F:126,277,378/Active site: His, Asp, Ser #status predicted

Query Match 35.6%; Score 127; DB 1; Length 433;
 Best Local Similarity 41.2%; Pred. No. 8,3e-07;
 Matches 28; Conservative 6; Mismatches 24; Indels 10; Gaps 3;

QY 1 CPMDSGLYREDQTSPPAGLRCLNWDAAQSGL-----ASAPVSGAGNHSYCRNPDED 52
 DB 72 CYQNGHSHYRKANRDLGRPCILAW-DSPTVLKMYAHMRSDALQGLGSKINCRNDP-N 129
 QY 53 PRGPMCYV 60
 DB 130 QRPMPCYV 137

RESULT 14
 US0600
 t-plasminogen activator (EC 3.4.21.68) gamma precursor - common vampire bat
 N/Alternate names: tissue plasminogen activator
 C/Species: Desmodus rotundus (common vampire bat)
 C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
 C/Accession: US0600
 R/Kraetschmar, J.; Haendler, B.; Langer, G.; Boldol, W.; Bringmann, P.; Alagon, A.; Dor
 Gene 105, 229-237, 1991
 A>Title: The plasminogen activator family from the salivary gland of the vampire bat Des
 A/Reference number: US0597; PMID:92039036; PMID:1937019
 A/Accession: US0600
 A/Molecule type: mRNA
 A/Residues: 1-394 <KRA>
 A/Cross-references: GB:M63990; NID:G16078; PIDN:AAA31595.1; PID:G16079
 A/Note: the authors translated the codon ATC for residue 75 as Thr
 C/Superfamily: tissue plasminogen activator; BGF homology; fibronectin type I repeat hom
 C/Keywords: fibrinolysis; glycoprotein; hydrolase; kringie; serine proteinase
 F:1-21/Domain: signal sequence #status predicted <PRO>
 F:22-36/Domain: propeptide #status predicted <PRO>
 F:37-394/Product: plasminogen activator gamma #status predicted <PLA>
 F:45-126/Domain: kringie homology <KR3>
 F:143-388/Domain: trypsin homology <TRY>
 F:45-126,66-108,97-121,131-262,174-190,182-251,276-351,308-324,341-369/Disulfide bonds:
 F:142-143/Cleavage site: His-Ser (plasmin) #status predicted
 F:189,238,345/Active site: His, Asp, Ser #status predicted
 F:315/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 35.3%; Score 126; DB 2; Length 394;
 Best Local Similarity 37.3%; Pred. No. 9,8e-07;
 Matches 25; Conservative 8; Mismatches 26; Indels 8; Gaps 2;

QY 1 CPMDSGLYREDQTSPPAGLRCLNWDAAQSGL-----LDAQSGLASAPVSGAGNHSYCRNPDED 53
 DB 45 CYQNGHSHYRKANRDLGRPCILAW-DSPTVLKMYAHMRSDALQGLGSKINCRNDP-N 104
 QY 54 PRGPMCYV 60
 DB 130 QRPMPCYV 137

DB 105 K-PMCYV 110

RESULT 15
 AA0332
 macrophage-stimulating protein 1 precursor - mouse
 N/Alternate names: hepatocyte growth factor-1-like protein
 C/Species: Mus musculus (house mouse)
 C/Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 18-Jun-1999
 A/Accession: AA0332; B40332
 R/Begen, S.J.F.; Stuart, L.A.; Han, S.; Jamison, C.S.
 Biochemistry 30, 9781-9791, 1991
 A>Title: Characterization of the mouse cDNA and gene coding for a hepatocyte growth fact
 A/Reference number: AA0332; MUID:92002017; PMID:1832957
 A/Accession: AA0332
 A/Molecule type: DNA
 A/Residues: 1-716 <DEG>
 A/Cross-references: GB:M74180; NID:G193831; PIDN:AAA50166.1; PID:G193832
 A/Accession: B40332
 A/Molecule type: mRNA
 A/Residues: 1-18, 'P', '20-716 <DEG2>
 A/Cross-references: GB:M74181; NID:G193833; PIDN:AAA50167.1; PID:G193834
 C/Genetics:
 A:Insertions: 18/1; 67/2; 105/1; 143/2; 189/1; 229/2; 269/1; 334/2; 378/1; 412/2; 458/1; 47
 C/Complex: disulfide-bonded heterodimer of chains derived from the same precursor
 C/Superfamily: hepatocyte growth factor; kringie homology; trypsin homology
 C/Keywords: duplication; glycoprotein; growth factor; kringie
 F:1-31/Domain: signal sequence #status predicted <SIG>
 F:19-488,489-716/Product: macrophage-stimulating protein 1 #status experimental <MAT>
 F:19-488,489-716/Product: alpha chain #status experimental <ACH>
 F:110-186/Domain: kringie homology <KR1>
 F:191-268/Domain: kringie homology <KR2>
 F:292-370/Domain: kringie homology <KR3>
 F:379-457/Domain: kringie homology <KR4>
 F:484-711/Domain: beta chain #status experimental <BCH>
 F:489-709/Domain: trypsin homology <TRY>
 F:72,173,305,620/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 35.3%; Score 126; DB 1; Length 716;
 Best Local Similarity 41.9%; Pred. No. 1,8e-06;
 Matches 26; Conservative 3; Mismatches 29; Indels 4; Gaps 2;

QY 1 CPMDSGLYREDQTSPPAGLRCLNWDAAQSGL-----LDAQSGLASAPVSGAGNHSYCRNPDEPRGPW 57
 DB 110 CYQNGHSHYRKANRDLGRPCILAW-DSPTVLKMYAHMRSDALQGLGSKINCRNDP-N 168
 QY 58 CY 59
 DB 169 CY 170

Search completed: April 7, 2003, 09:19:20
 Job time : 5.86486 secs

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:24 ; Search time 2.74428 Seconds
(without alignments)
906.823 Million cell updates/sec

Title: US-10-057-951-2_COPY_25_84
Perfect score: 357
Sequence: 1 CFWDNGHLYREDQTSPPAGL.....GNHSYCRNPDEPRGFWCYV 60

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	# Query Match	Length DB	ID	Description
1	154	43.1	566	1	TPA_BOVIN
2	146	40.9	562	1	TPA_HUMAN
3	146	40.9	653	1	HGFA_MOUSE
4	141	39.5	559	1	TPA_RAT
5	141	39.5	655	1	HGFA_HUMAN
6	136	38.1	559	1	TPA_MOUSE
7	134	37.5	431	1	URTB_DESRO
8	132	37.0	431	1	URTB_DESRO
9	131	36.7	431	1	UROK_PAPCY
10	131	36.7	431	1	UROK_HUMAN
11	130	36.4	442	1	UROK_PIG
12	127	35.6	433	1	UROK_BOVIN
13	126	35.3	394	1	URTG_DESRO
14	126	35.3	716	1	HGFL_MOUSE
15	123	34.5	169	1	PLMN_RAT
16	121	33.9	810	1	PLMN_HUMAN
17	119	33.3	433	1	UROK_MOUSE
18	119	33.3	728	1	HGF_HUMAN
19	118	33.1	432	1	UROK_RAT
20	118	33.1	728	1	HGF_MOUSE
21	118	33.1	810	1	PLMN_ERIEU
22	115	32.2	477	1	URTI_DESRO
23	113.5	31.8	333	1	PLMN_CANPA
24	113.5	31.8	790	1	PLMN_PIG
25	113.5	31.8	4548	1	APOA_HUMAN
26	113	31.7	810	1	PLMN_MACMU
27	112	31.4	728	1	HGF_RAT
28	111.5	31.2	603	1	FA12_CAVPO
29	111	31.1	812	1	PLMN_BOVIN
30	110	30.8	473	1	KREM_MOUSE
31	110	30.8	473	1	KREM_RAT
32	110	30.8	475	1	KREM_HUMAN
33	109	30.5	711	1	HGFL_HUMAN

34	108	30.3	625	1	THRB_BOVIN	P00735 bos taurus
35	106.5	29.8	434	1	UROK_CHICK	P15120 gallus galli
36	105	29.4	622	1	THRB_HUMAN	P00734 homo sapien
37	104.5	29.3	812	1	PLMN_MOUSE	P20918 mus musculus
38	98.5	27.6	593	1	FA12_BOVIN	P08140 bos taurus
39	97.5	27.3	343	1	PLMN_SHEEP	P01286 ovis aries
40	97	27.2	617	1	THRB_RAT	P18292 ratius norv
41	95.5	26.8	338	1	PLMN_HORSE	P00010 equus cabal
42	95	26.6	618	1	THRB_MOUSE	P19221 mus musculus
43	94.5	26.5	325	1	PLMN_PETMA	P33574 petromyzon
44	94.5	26.5	615	1	FA12_HUMAN	P00748 homo sapien
45	93.5	26.2	1420	1	APOA_MACMU	P14417 macaca mula

ALIGNMENTS

```

RESULT 1
TPA_BOVIN          STANDARD;          PRT;          566 AA.
AC  Q28198;
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE  (t-PA) (t-plasminogen activator).
GN  PLAT.
OS  Bos taurus (Bovine).
OC  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC  Bovidae; Bovinae; Bos.
OX  NCBI_Taxid=9913;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Kidney;
RA  Ravn P., Berglund L., Petersen T.E.;
RT  "Cloning and characterization of the bovine plasminogen activators uPA
RT  and tPA.";
RL  Int. Dairy J. 5:605-617(1995).
CC  -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN
CC  TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC  CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC  ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC  MANY OTHER PHYSIOLOGICAL EVENTS.
CC  -!- CATALYTIC ACTIVITY: Specific cleavage of Arg|-Val bond in
CC  plasminogen to form plasmin.
CC  -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC  BOND.
CC  -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC  -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC  PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC  ARG-314 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC  -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC  CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC  -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC  -!- SIMILARITY: CONTRAINS 1 EGF-LIKE DOMAIN.
CC  -!- SIMILARITY: CONTRAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC  -!- SIMILARITY: CONTRAINS 2 KRINGLE DOMAINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (see http://www.isb.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X85800; CAA59795.1; -.
DR  HSSP; P00750; 1RTP.
DR  MEROPS; S01.232; -.
DR  InterPro; IPR001314; Chymotrypsin.
DR  InterPro; IPR000561; EGF-like.
DR  InterPro; IPR000083; Fibrinctn1.

```

```
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; Kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR Prodom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Plasma; Kringie; EGF-like domain; Repeat; Signal.
KW Plasmogen activation; Hydrolyase; Serine protease; Glycoprotein;
FT SIGNAL 1 21 BY SIMILARITY.
FT PROPEP 22 33 TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 34 566 TISSUE-TYPE PLASMINOGEN ACTIVATOR A
FT CHAIN 34 314 TISSUE-TYPE PLASMINOGEN ACTIVATOR B
FT CHAIN 315 566 TISSUE-TYPE PLASMINOGEN ACTIVATOR B
FT CHAIN 315 566 FIBRONECTIN TYPE-I.
FT DOMAIN 40 82 EGF-LIKE.
FT DOMAIN 83 121 KRINGLE 1.
FT DOMAIN 128 209 KRINGLE 2.
FT DOMAIN 219 300 SERINE PROTEASE.
FT DOMAIN 315 566 CHARGE RELAY SYSTEM.
FT ACT_SITE 361 410 CHARGE RELAY SYSTEM.
FT ACT_SITE 410 410 CHARGE RELAY SYSTEM.
FT ACT_SITE 410 410 CHARGE RELAY SYSTEM.
FT ACT_SITE 517 517 CHARGE RELAY SYSTEM.
FT DISULFID 42 72 BY SIMILARITY.
FT DISULFID 70 79 BY SIMILARITY.
FT DISULFID 87 98 BY SIMILARITY.
FT DISULFID 92 109 BY SIMILARITY.
FT DISULFID 111 120 BY SIMILARITY.
FT DISULFID 128 209 BY SIMILARITY.
FT DISULFID 149 191 BY SIMILARITY.
FT DISULFID 180 204 BY SIMILARITY.
FT DISULFID 219 300 BY SIMILARITY.
FT DISULFID 240 282 BY SIMILARITY.
FT DISULFID 271 295 BY SIMILARITY.
FT DISULFID 303 434 INTERCHAIN (BY SIMILARITY).
FT DISULFID 346 362 BY SIMILARITY.
FT DISULFID 354 423 BY SIMILARITY.
FT DISULFID 448 523 BY SIMILARITY.
FT DISULFID 480 496 BY SIMILARITY.
FT CARBOHYD 513 541 BY SIMILARITY.
FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 487 487 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 566 AA; 63701 MW; 2EB6BB4E32276C3 CRC64;
```

```
Query Match 43.1%; Score 154; DB 1; Length 566;
Best Local Similarity 44.9%; Pred. No. 2e-10;
Matches 31; Conservative 6; Mismatches 20; Indels 12; Gaps 3;
```

```
QY 1 CWDNNGHLREDOQSPAPGLRCLNWDIAQSGIASAPV-----GAGNHSYCRNPDE 51
DB 128 CYKDGVAHRYSTSTSGAECANW--NSSGLAMKPYSGRRPNAIRLIGHNHYCRNPDQ 185
QY 52 DRGPKVCYV 60
DB 186 DSK-PWCYV 193
```

```
TPA_HUMAN
ID TPA_HUMAN STANDARD; PRT; 562 AA.
AC P00750; Q15103;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator) (Alteplase) (Retepase).
GN PLAT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=83115262; PubMed=6337343;
RA Pennica D., Holmes W.E., Kohr W.J., Harkins R.N., Vehar G.A.,
RA Ward C.A., Bennett W.F., Velterton E., Seeburg P.H., Heyneker H.L.,
RT Goeddel D.V., Collen D.;
RT "Cloning and expression of human tissue-type plasminogen activator
RT cDNA in E. coli.";
RL Nature 301:214-221(1983).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal lung;
RX MEDLINE=88262579; PubMed=3133640;
RA Sasaki H., Saito Y., Hayashi M., Otsuka K., Nima M.;
RT "Nucleotide sequence of the tissue-type plasminogen activator cDNA
RT from human fetal lung cells.";
RL Nucleic Acids Res. 16:5695-5695(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88054470; PubMed=2824147;
RA Reddy V.B., Garramone A.J., Sasak H., Wei C.-M., Watkins P., Galli J.,
RA Heising N.;
RT "Expression of human uterine tissue-type plasminogen activator in
RT mouse cells using BPV vectors.";
RL DNA 6:461-472(1987).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=86196143; PubMed=3009482;
RA Friesner Degen S.J., Rajput B., Reich E.;
RT "The human tissue plasminogen activator gene.";
RL J Biol. Chem. 261:6972-6985(1986).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=84298137; PubMed=6089198;
RA NY T., Elgh F., Lund B.;
RT "The structure of the human tissue-type plasminogen activator gene:
RT correlation of intron and exon structures to functional and
RT structural domains.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359(1984).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=86284200; PubMed=3090401;
RA Harris T.J., Patel T., Marston F.A., Little S., Emrage J.S.;
RA Opdenakker G., Volckaert G., Rombauts W., Billiau A., Somer P.;
RT "Cloning of cDNA coding for human tissue-type plasminogen activator
RT and its expression in Escherichia coli.";
RL Mol. Biol. Med. 3:279-292(1986).
RN [7]
RP SEQUENCE OF 212-361 FROM N.A.
RX MEDLINE=83169656; PubMed=6572897;
RA Edlund T., Ny T., Raanby M., Heden L.-O., Palm G., Holmgren E.,
RA Josephson S.;
RT "Isolation of cDNA sequences coding for a part of human tissue
RT plasminogen activator.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:349-352(1983).
RN [8]
RP SEQUENCE OF 1-36 FROM N.A.
RX MEDLINE=85289318; PubMed=3161893;
RA Fisher R., Waller E.K., Grossi G., Thompson D., Tizard R.,
```

RA Schleuning W.-D.;
 RT "Isolation and characterization of the human tissue-type plasminogen
 RT activator structural gene including its 5' flanking region.";
 RL J. Biol. Chem. 260:11223-11230 (1985).
 RN [9]
 RP SEQUENCE OF 31-562 FROM N.A.
 RX MEDLINE=91291340; PubMed=1366881;
 RA Itagaki Y., Yasuda H., Morinaga T., Mitsuda S., Higashio K.;
 RT "Purification and characterization of tissue plasminogen activator
 RT secreted by human embryonic lung diploid fibroblasts, IMR-90 cells.";
 RL Agric. Biol. Chem. 55:1225-1232 (1991).
 RN [10]
 RP SEQUENCE OF 36-562.
 RC TISSUE-Melanoma;
 RX MEDLINE=85000468; PubMed=64333976;
 RA Pohl G., Kaellstroem M., Bergsdorf N., Wallen P., Joernvall H.;
 RT "Tissue plasminogen activator: peptide analyses confirm an indirectly
 RT derived amino acid sequence, identify the active site serine residue,
 RT establish glycosylation sites, and localize variant differences.";
 RL Biochemistry 23:3701-3707 (1984).
 RN [11]
 RP SEQUENCE OF 33-52 AND 311-330.
 RC TISSUE-Melanoma;
 RX MEDLINE=83209620; PubMed=6682760;
 RA Wallen P., Pohl G., Bergsdorf N., Raanby M., Ny T., Joernvall H.;
 RT "Purification and characterization of a melanoma cell plasminogen
 RT activator.";
 RL Eur. J. Biochem. 132:681-686 (1983).
 RN [12]
 RP SEQUENCE FROM N.A. (SHORT ISOFORM).
 RC TISSUE=umbilical vein;
 RX MEDLINE=90192129; PubMed=2107528;
 RA Siebert P.D., Fong K.;
 RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from
 RT human endothelial cells.";
 RL Nucleic Acids Res. 18:1086-1086 (1990).
 RN [13]
 RP STRUCTURE OF CARBOHYDRATES.
 RX MEDLINE=90092112; PubMed=2513186;
 RA Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.;
 RT "Carbohydrate structure of recombinant human uterine tissue
 RT plasminogen activator expressed in mouse epithelial cells.";
 RL Eur. J. Biochem. 186:273-286 (1989).
 RN [14]
 RP CARBOHYDRATE-LINKAGE SITE THR-96.
 RX MEDLINE=91159408; PubMed=1900431;
 RA Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.;
 RT "Tissue plasminogen activator has an O-linked fucose attached to
 RT chreonine-61 in the epidermal growth factor domain.";
 RL Biochemistry 30:2311-2314 (1991).
 RN [15]
 RP DISULFIDE BONDS IN KRINGLE 2.
 RX MEDLINE=91244765; PubMed=1645336;
 RA Vlahos C.J., Wilhelm O.G., Hassell T., Taskunas S.R., Bang N.U.;
 RT "Disulfide pairing of the recombinant kringle-2 domain of tissue
 RT plasminogen activator produced in *Bescherichia coli*.";
 RL J. Biol. Chem. 266:10070-10072 (1991).
 RN [16]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF CATALYTIC DOMAIN.
 RX MEDLINE=96200985; PubMed=8613982;
 RA Lamba D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,
 RA Bode W.;
 RT "The 2.3 Å crystal structure of the catalytic domain of recombinant
 RT two-chain human tissue-type plasminogen activator.";
 RL J. Mol. Biol. 258:117-135 (1996).
 RN [17]
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.
 RX MEDLINE=97449126; PubMed=9305622;
 RA Renatus M., Engh R.A., Stubbs M.T., Huber R., Fischer S., Kohnert U.,
 RA Bode W.;
 RT "Lysine 156 promotes the anomalous proenzyme activity of tPA: X-ray
 RT crystal structure of single-chain human tPA.";
 RL EMBO J. 16:4797-4805 (1997).

RN [18]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF KRINGLE 2.
 RX MEDLINE=92118803; PubMed=1310033;
 RA de Vos A., Ulsch M.H., Kelley R.F., Padmanabhan K., Tulinsky A.,
 RA Westbrook M.L., Kosiakof A.A.;
 RT "Crystal structure of the kringle 2 domain of tissue plasminogen
 RT activator at 2.4-Å resolution.";
 RL Biochemistry 31:270-279 (1992).
 RN [19]
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=90122799; PubMed=2558718;
 RA Byeon I.-J.L., Kelley R.F., Lillias M.;
 RT "1H NMR structural characterization of a recombinant kringle 2 domain
 RT from human tissue-type plasminogen activator.";
 RL Biochemistry 28:9350-9360 (1989).
 RN [20]
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=91200042; PubMed=1901789;
 RA Byeon I.-J.L., Kelley R.F., Lillias M.;
 RT "Kringle-2 domain of the tissue-type plasminogen activator. 1H-NMR
 RT assignments and secondary structure.";
 RL Eur. J. Biochem. 197:155-165 (1991).
 RN [21]
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=92106329; PubMed=1762144;
 RA Byeon I.-J.L., Lillias M.;
 RT "Solution structure of the tissue-type plasminogen activator kringle
 RT 2 domain complexed to 6-aminohexanoic acid an antifibrinolytic
 RT drug.";
 RL J. Mol. Biol. 222:1035-1051 (1991).
 RN [22]
 RP STRUCTURE BY NMR OF 38-85.
 RX MEDLINE=92292163; PubMed=1602484;
 RA Downing A.K., Driscoll P.C., Harvey T.S., Dudgeon T.J., Smith B.O.,
 RA Baron M., Campbell I.D.;
 RT "Solution structure of the fibrin binding finger domain of
 RT tissue-type plasminogen activator determined by 1H nuclear magnetic
 RT resonance.";
 RL J. Mol. Biol. 225:821-833 (1992).
 RN [23]
 RP STRUCTURE BY NMR OF 36-126.
 RX MEDLINE=96027104; PubMed=7582899;
 RA Smith B.O., Downing A.K., Driscoll P.C., Dudgeon T.J., Campbell I.D.;
 RT "The solution structure and backbone dynamics of the fibronectin type
 RT I and epidermal growth factor-like pair of modules of tissue-type
 RT plasminogen activator.";
 RL Structure 3:823-833 (1995).
 CC -1- FUNCTION: CONVERGES THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN
 CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
 CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
 CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
 CC MANY OTHER PHYSIOLOGICAL EVENTS.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: BINDS TO FIBRIN WITH HIGH AFFINITY. THIS INTERACTION

Query Match 40.9%; Score 146; DB 1; Length 562;
 Best Local Similarity 43.5%; Pred. No. 1.6e-09;
 Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

QY 1 CPWNGHLYREDQSPAPRGCLMWDAOGSLASAPV-----GAGNHSYCRPDE 51
 DB 127 CYEDGDSYRTGTWSTVSSGAEACTW--NSSALAQKPYSGRRPDAIRLGLGNHNYCRPPR 184
 QY 52 DPRGWCYV 60
 DB 185 DSK-PWCYV 192

RESULT 3
 HGFA_MOUSE STANDARD; PRT; 653 AA.
 ID HGFA_MOUSE
 AC Q9R098; Q9JXV4;

```

DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF
DE activator) (HGFA).
GN HGFAc.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN (1)
RP SEQUENCE FROM N.A.
RA STRAIN=BALB/c;
RA Itoh H., Kataoka H., Koono H.;
RT "Mouse hepatocyte growth factor activator.";
RL Submitted (MAR-1993) to the EMBL/GenBank/DBD databases.
RN (2)
RP SEQUENCE FROM N.A.
RA van Adelsberg J.S., Sehgal S., Kukes A., Brady C., Barasch J.,
RA Yang J., Han Y.;
RT "Activation of HGF by endogenous HGF activator is required for
RT mesenchymal kidney morphogenesis in vitro.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBD databases.
CC -1- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY CONVERTING
CC IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM (BY SIMILARITY).
CC -1- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A
CC DISULFIDE BOND (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN
CC PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE II DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as their content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF099017; AAF02489.1; -
DR EMBL; AF224724; AAF34712.1; -
DR HSSP; P00763; IDPO.
DR MEROPS; S01.228; -.
DR MGD; MGI:1859281; Hgfac.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF-2.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR000083; Fibronectn.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00068; EGF_2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; ENTPERTI.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR ProDom; PD000995; FN_Type_II; 1.
DR SMART; SM00181; EGF_2.
DR SMART; SM00058; fn1; 1.
DR SMART; SM00059; fn2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.

```

DR	PROSITE	PS01253; FIBRONECTIN_1; 1.
DR <td>PROSITE</td> <td>PS00023; FIBRONECTIN_2; 1.</td>	PROSITE	PS00023; FIBRONECTIN_2; 1.
DR	PROSITE	PS00021; KRINGLE_1; 1.
DR	PROSITE	PS00070; KRINGLE_2; 1.
DR	PROSITE	PS0240; TRYPSIN_DOM; 1.
DR	PROSITE	PS00134; TRYPSIN_HIS; 1.
DR	PROSITE	PS00135; TRYPSIN_SER; 1.
KW	Hydrolase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;	
KM	Egr-like domain; Repeat; Zymogen.	
FT	SIGNAL	1 29
FT	PROPE	30 369
FT	CHAIN	370 405
FT	CHAIN	406 653
FT	DOMAIN	105 145
FT	DOMAIN	157 195
FT	DOMAIN	197 237
FT	DOMAIN	238 276
FT	DOMAIN	283 364
FT	DOMAIN	406 653
FT	ACT_SITE	445 445
FT	ACT_SITE	445 445
FT	ACT_SITE	556 556
FT	ACT_SITE	556 596
FT	ACT_SITE	105 130
FT	DISULFID	119 145
FT	DISULFID	161 172
FT	DISULFID	166 183
FT	DISULFID	185 194
FT	DISULFID	199 227
FT	DISULFID	225 234
FT	DISULFID	242 253
FT	DISULFID	247 264
FT	DISULFID	266 275
FT	DISULFID	283 364
FT	DISULFID	304 346
FT	DISULFID	335 359
FT	DISULFID	392 519
FT	DISULFID	430 446
FT	DISULFID	438 508
FT	DISULFID	533 602
FT	DISULFID	565 581
FT	DISULFID	592 620
FT	CARBOHYD	39 39
FT	CARBOHYD	47 47
FT	CARBOHYD	63 63
FT	CARBOHYD	287 287
FT	CARBOHYD	466 466
FT	CARBOHYD	544 544
FT	CONFLICT	164 164
SC	SEQUENCE	653 AA; 70567 MW; 88B4B2055D7FDC CXC64;
Query Match		
Best Local Similarity	40.9%;	Score 146; DB 1; Length 653;
Matches 32; Conservative	47.1%;	Pred. No. 1.9e-09;
	7; Mismatches	19; Indels 10; Gaps 3;
Qy	1	CPWMDGHLYREDQSPAPGLRCLNN-----LDAQGLASAPVSGAGNSYCNRPED 52
Db	283	CFLGNGTGYRGVASTAAGSLCLANWSDLYOEHLVDS-VAAAVLLIGLGPAPHYCRNPDKD 341
Qy	53	PRGPWCYV 60
Db	342	ER-PWCYV 348
RESULT 4		
TPA_RAT	TPA_RAT	STANDARD; PRT; 559 AA.
AC	P19637;	
DT	01-FEB-1991	(Rel. 17, Created)
DT	01-NOV-1991	(Rel. 20, Last sequence update)
DT	15-JUN-2002	(Rel. 41, Last annotation update)
DE	Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)	

ID HGFA HUMAN STANDARD; PRT; 655 AA.
AC 004756; 014726;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF activator) (HGFA).
GN HGFA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC TISSUE=Liver, and Serum;
RX MEDLINE=93252878; PubMed=763665;
RA Miyazawa K., Shimomura T., Kitamura A., Kondo J., Morimoto Y., Kitamura N.;
RT "Molecular cloning and sequence analysis of the cDNA for a human serine protease responsible for activation of hepatocyte growth factor. Structural similarity of the protease precursor to blood coagulation factor XII.";
RT J. Biol. Chem. 268:10024-10028 (1993).
RN [2]
RP SEQUENCE OF 40-655 FROM N.A.
RA Zhao S., Ode11 C.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY CONVERTING IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM.
CC -1- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A DISULFIDE BOND.
CC -1- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM.
CC -1- TISSUE SPECIFICITY: LIVER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE II DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 IS THE INITIATOR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D14012; BAA03113.1; -;
DR EMBL; Z69923; CA93803.1; -;
DR PIR; A46688; A46688.
DR HSRP; P00763; IDPO.
DR MEROPS; S01.228; -;
DR GeneW; HGNC:4894; HGFA.
DR MIM; 604552; -;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR000562; FN Type II.
DR InterPro; IPR000083; p130catl.
DR InterPro; IPR000001; kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00039; fn2; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FMTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; kringle; 1.
DR ProDom; PD000995; FN_Type_II; 1.

DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR HydroLase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
KW EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 30
FT PROPEP 31 372
FT CHAIN 373 407
FT CHAIN 408 655
FT CHAIN 655 655
FT DOMAIN 108 148
FT DOMAIN 160 198
FT DOMAIN 200 240
FT DOMAIN 241 279
FT DOMAIN 286 367
FT DOMAIN 408 655
FT ACT_SITE 447 447
FT ACT_SITE 497 497
FT ACT_SITE 598 598
FT ACT_SITE 108 133
FT DISULFID 122 148
FT DISULFID 164 175
FT DISULFID 199 216
FT DISULFID 188 197
FT DISULFID 202 230
FT DISULFID 228 237
FT DISULFID 245 256
FT DISULFID 250 267
FT DISULFID 269 278
FT DISULFID 286 367
FT DISULFID 307 349
FT DISULFID 338 362
FT DISULFID 394 521
FT DISULFID 432 448
FT DISULFID 440 510
FT DISULFID 535 604
FT DISULFID 567 583
FT DISULFID 594 632
FT CARBOHYD 448 46
FT CARBOHYD 290 230
FT CARBOHYD 468 468
FT CARBOHYD 492 492
FT CARBOHYD 546 546
FT CONFLICT 644 644
SQ SEQUENCE 655 AA; 70681 MW; 2CF72P1E1B862ED7 CRC64;
Query Match 39.5%; Score 141; DB 1; Length 655;
Best Local Similarity 45.6%; Pred. No. 7.3e-05;
Matches 31; Conservative 6; Mismatches 21; Indels 10; Gaps 3;
QY 1 CFWDNGHLYREDTSPAPGLRLCNM-----LDAOSGLASAPVSGAGNHSYCRNPDD 52
DB 286 CFLNGTGGRVASTSASGLSCIANNSDLVQELHVDV-VGAALLGLGPHAYCRNPDD 344
QY 53 PRGPWCYV 60
DB 345 ER-PMCTV 351
RESULT 6
TPA_MOUSE

TPA MOUSE STANDARD; PRT; 559 AA.
 AC P11214;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tpa)
 DE (t-PA) (t-plasminogen activator).
 CN PLAT.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N. A.
 RX MEDLINE=88087303; PubMed=2826484;
 RA Rickles R.J., Darrow A.L., Strickland S.;
 RT "Molecular cloning of complementary DNA to mouse tissue plasminogen
 RT activator mRNA and its expression during F9 teratocarcinoma cell
 RT differentiation";
 RT J. Biol. Chem. 263:1563-1569(1988).
 RU J. Biol. Chem. 263:1563-1569(1988).
 CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN
 CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
 CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
 CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
 CC MANY OTHER PHYSIOPATHOLOGICAL EVENTS.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
 CC BOND.
 CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
 CC -1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
 CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
 CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
 CC -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
 CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE 1 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; J03520; AAA0470.1; -.
 CC PIR; A29941; A29941.
 CC HSSP; P00750; 1A5H.
 CC MEROPS; S01.232; -.
 CC MGD; MGI:97610; Plac.
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR000561; EGF-like.
 CC InterPro; IPR000083; Fibrinctn1.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR001254; Ser_protease_Try.
 CC Pfam; PF00008; EGF_1.
 CC Pfam; PF00039; Efn_1.
 CC Pfam; PF00051; Kringle_2.
 CC Pfam; PF00089; trypsin_1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00018; KRINGLE.
 CC PRODOM; PD000395; Kringle_2.
 CC SMART; SM00181; EGF_1.
 CC SMART; SM00058; EFN_1.
 CC SMART; SM00130; KR_2.
 CC SMART; SM00020; TRY_PSP_1.
 CC PROSITE; PS00022; EGF_1; 1.
 CC PROSITE; PS01186; EGF_2; 1.
 CC PROSITE; PS01253; FIBRONECTIN_1; 1.
 CC PROSITE; PS00021; KRINGLE_1; 2.

DR PROSITE; PS00070; KRINGLE_2; 2.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SRR; 1.
 KW plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
 KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
 FT SIGNAL 1 17
 FT PROPEP 18 29
 FT CHAIN 30 559
 FT CHAIN 30 306
 FT CHAIN 309 559
 FT CHAIN 309 559
 FT DOMAIN 36 78
 FT DOMAIN 79 117
 FT DOMAIN 124 205
 FT DOMAIN 213 294
 FT DOMAIN 309 559
 FT ACT_SITE 355 355
 FT ACT_SITE 404 404
 FT ACT_SITE 510 510
 FT DISULFID 38 68
 FT DISULFID 66 75
 FT DISULFID 83 94
 FT DISULFID 88 105
 FT DISULFID 107 116
 FT DISULFID 124 205
 FT DISULFID 145 187
 FT DISULFID 176 200
 FT DISULFID 213 294
 FT DISULFID 234 276
 FT DISULFID 265 289
 FT DISULFID 297 428
 FT DISULFID 340 356
 FT DISULFID 348 417
 FT DISULFID 442 516
 FT DISULFID 474 490
 FT DISULFID 506 534
 FT CARBOHYD 149 149
 FT CARBOHYD 481 481
 FT SEQUENCE 559 AA; 63110 MW; 4ACE57DCE6A282A5 CRC64;
 Query Match 38.1%; Score 136; DB 1; Length 559;
 Best Local Similarity 40.6%; Pred. No. 2,4e-08;
 Matches 28; Conservative 8; Mismatches 21; Indels 12; Gaps 3;
 QY 1 CFWMGNHLYREDQSPAPGRCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 51
 DB 124 CFEBOGITYRTGTWSTASGACINW--NSSVLSIKPYNARRPNAIKLGLGNHNYCRNPDR 181
 QY 52 DPRGPWCYV 60
 DB 182 DLK-PWCYV 189
 RESULT 7
 ID URTB DESRO STANDARD; PRT; 431 AA.
 AC P98121;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Salivary plasminogen activator beta precursor (EC 3.4.21.68) (DSPA
 DE beta).
 OS Desmodus rotundus (Vampire bat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
 OC Desmodontinae; Desmodus.
 OX NCBI_Taxid=9430;
 RN [1]
 RP SEQUENCE FROM N. A.
 RC TISSUE=Salivary gland;
 RX MEDLINE=92039036; PubMed=1937019;

RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
 RA Alagon A., Donner P., Schleuning W.D.;
 RT "The plasminogen activator family from the salivary gland of the
 RT vampire bat *Desmodus rotundus*: cloning and expression.";
 RL Gene 105:229-237(1991).
 RN (2)
 RP CHARACTERIZATION.
 RX MEDLINE=93393059; PubMed=1309059;
 RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
 RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Wilt W.,
 RA Donner P.;
 RT "Plasminogen activators from the saliva of *Desmodus rotundus* (common
 RT vampire bat): unique fibrin specificity.";
 RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
 CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
 CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
 CC AGENT.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).

 DR EMBL: M63989; AAA31594.1; -
 DR HSSP: P98119; 1A51.
 DR MEROPS: S01.239; -
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam: PF00008; EGF_1.
 DR Pfam: PF00051; kringle_1.
 DR Pfam: PF00089; trypsin_1.
 DR Pfam: PF00722; CHYMOTRYPSIN.
 DR PRINTS: PR00018; KRINGLE.
 DR ProDom: PD000395; Kringle_1.
 DR SMART: SM00181; EGF_1.
 DR SMART: SM00130; KR_1.
 DR SMART: SM00020; Tryp_Spec_1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS00021; KRINGLE_1; 1.
 DR PROSITE: PS00070; KRINGLE_2; 1.
 DR PROSITE: PS02040; TRYPsin_DOM_1.
 DR PROSITE: PS00134; TRYPsin_HIS_1.
 DR PROSITE: PS00135; TRYPsin_SER_1.
 KW Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
 KW Signal; EGF-like domain; Signal; Multigene family.
 FT SIGNAL 1 36
 FT CHAIN 37 431
 FT DOMAIN 82 163
 FT DOMAIN 179 431
 FT ACT_SITE 226 275
 FT ACT_SITE 275 275
 FT ACT_SITE 382 382
 FT DISULFID 41 52
 FT DISULFID 46 63
 FT DISULFID 65 74
 FT DISULFID 82 163
 FT DISULFID 103 145
 FT DISULFID 134 158
 FT DISULFID 168 299
 FT DISULFID 211 227
 BY SIMILARITY.

FT DISULFID 219 288 BY SIMILARITY.
 FT DISULFID 313 388 BY SIMILARITY.
 FT DISULFID 345 361 BY SIMILARITY.
 FT DISULFID 378 406 BY SIMILARITY.
 FT CARBOHYD 139 139 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 352 352 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 431 AA; 48221 MW; 699B5E75B162CBF CRC64;
 Query Match 37.5%; Score 134; DB 1; Length 431;
 Best Local Similarity 38.6%; Pred. No. 3.1e-08;
 Matches 27; Conservative 9; Mismatches 20; Indels 14; Gaps 3;
 QY 1 CFPDNGHILHREDTSPAPGRLCINWIDAQSG-----ASAPVSGAGHSTYCRNP 50
 DB 82 CYDDQGVYKGTWSTSESGAQCINM---NSNLTRRTYNGRSDAITLGLGNHNYCRNP 138
 QY 51 EDRGPMPCVY 60
 DB 139 NNSK-PWCIV 147
 RESULT 8
 ID URT2 DESRO STANDARD: PRT; 477 AA.
 AC P15638;
 DT 01-APR-1990 (Rel. 14; Created)
 DT 01-FEB-1996 (Rel. 33; Last sequence update)
 DT 15-JUN-2002 (Rel. 41; Last annotation update)
 DE Salivary plasminogen activator alpha 2 precursor (EC 3.4.21.68) (DSBA
 OS alpha-2) (BART-PA) (T-plasminogen activator).
 OS *Desmodus rotundus* (Vampire bat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
 OC Desmodontinae; *Desmodus*.
 OC NCBI_Taxid=9430;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=salivary gland;
 RX MEDLINE=92039036; PubMed=1937019;
 RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
 RA Alagon A., Donner P., Schleuning W.D.;
 RT "The plasminogen activator family from the salivary gland of the
 RT vampire bat *Desmodus rotundus*: cloning and expression.";
 RL Gene 105:229-237(1991).
 RN [2]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=salivary gland;
 RX MEDLINE=90036867; PubMed=2509450;
 RA Gardell S.J., Duong L.T., Diehl R.E., York J.D., Hare T.R.,
 RA Register R.B., Jacobs J.W., Dixon R.A.F., Friedman P.A.;
 RT "Isolation, characterization, and cDNA cloning of a vampire bat
 RT salivary plasminogen activator.";
 RL J. Biol. Chem. 264:17947-17952(1989).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=93393059; PubMed=1309059;
 RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
 RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Wilt W.,
 RA Donner P.;
 RT "Plasminogen activators from the saliva of *Desmodus rotundus* (common
 RT vampire bat): unique fibrin specificity.";
 RL Ann. N.Y. Acad. Sci. 667:395-403(1992)
 CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
 CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
 CC AGENT.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- ENZYME REGULATION: ACTIVITY TOWARD PLASMINOGEN IS STIMULATED IN
 CC THE PRESENCE OF FIBRIN I.
 CC -1- SUBUNIT: MONOMER.
 CC -1- DOMAIN: THE FIBRINECTIN TYPE-I DOMAIN MEDIATES BINDING TO FIBRIN,
 CC AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED
 CC STIMULATION OF ACTIVITY.

DR PROSITE; PS50070; KRINGLE 2; 1.
 DR PROSITE; PS50240; TRYPSIN DOM; 1.
 DR PROSITE; PS00134; TRYPSIN HIS; 1.
 DR PROSITE; PS00135; TRYPSIN SER; 1.
 KM Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
 KM Kringle; EGF-like domain; zymogen; signal.
 FT SIGNAL 1 20
 FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 FT CHAIN 21 176 CHAIN A (BY SIMILARITY).
 FT CHAIN 155 176 SHORT A CHAIN (A1) (BY SIMILARITY).
 FT CHAIN 178 433 CHAIN B (BY SIMILARITY).
 FT DOMAIN 26 62 EGF-LIKE.
 FT DOMAIN 69 150 KRINGLE.
 FT DOMAIN 151 177 CONNECTING PEPTIDE.
 FT DOMAIN 178 433 SERINE PROTEASE.
 FT DISULFID 30 38 BY SIMILARITY.
 FT DISULFID 32 50 BY SIMILARITY.
 FT DISULFID 52 61 BY SIMILARITY.
 FT DISULFID 167 298 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 208 224 BY SIMILARITY.
 FT DISULFID 216 287 BY SIMILARITY.
 FT DISULFID 315 384 BY SIMILARITY.
 FT DISULFID 347 363 BY SIMILARITY.
 FT DISULFID 374 402 BY SIMILARITY.
 FT ACT_SITE 223 223 CHARGE RELAY SYSTEM.
 FT ACT_SITE 274 274 CHARGE RELAY SYSTEM.
 FT ACT_SITE 378 378 CHARGE RELAY SYSTEM.
 FT CARBOHYD 324 324 N-LINKED (GLCNAC...) (BY SIMILARITY).
 SQ SEQUENCE 433 AA; 48595 MW; 816D22DFEDDC8792 CRC64;

Query Match 37.0%; Score 132; DB 1; Length 433;
 Best local similarity 40.3%; Pred. No. 5.3e-08;
 Matches 27; Conservative 7; Mismatches 25; Indels 8; Gaps 2;

QY 1 CFWDNGHLYREDQTSPPAPGRLWLDL-----QSGLSAPVSGAGNHYCRNEDDP 53
 DB 69 CYEGNCHFYRGKASDTWGRSCLAMNSATVYQQTTHAHRSDALQLGKHYCNRPD-NR 127

QY 54 RGPWCYV 60
 DB 128 RRMWCYV 134

RESULT 10
 ID UROK_HUMAN STANDARD; PRT; 431 AA.
 AC P00749; O15844; O16188; O969W6;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 20-MAR-1987 (Rel. 04, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)
 DE (U-plasminogen activator).
 GN PLAU.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85215647; PubMed=2987867;
 RA Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Blaai F.,
 RT "The human urokinase-plasminogen activator gene and its promoter."
 RL Nucleic Acids Res. 13:2759-2771(1985).
 RN (2)
 RP SEQUENCE FROM N.A.
 RA Holmes W.E., Pennic D., Blaber M., Rey M.W., Guenzler W.A.,
 RT Steffens G.J., Heynaker H.L.;
 RT "Cloning and expression of the gene for pro-urokinase in Escherichia
 coli.";
 RL Biotechnology 3:923-929(1985).
 RN (3)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85056954; PubMed=2415429;

RA Nagai M., Hiramatsu R., Kaneda T., Hayasuke N., Atimura H.,
 RA Nishida M., Suyama T.;
 RT "Molecular cloning of cDNA coding for human preprourokinase."
 RL Gene 36:183-188(1985).
 RN (4)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85203359; PubMed=3888571;
 RA Jacobs P., Cravador A., Lortiau R., Brockly F., Colau B., Chuchana P.,
 RA van Elsen A., Herzog A., Boilen A.;
 RT "Molecular cloning, sequencing, and expression in Escherichia coli of
 human preprourokinase cDNA."
 RL DNA 4:139-146(1985).
 RN (5)
 RP SEQUENCE FROM N.A.
 RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
 RA Nickerson D.A.;
 RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN (6)
 RP SEQUENCE FROM N.A.
 RC Tissue=Lung;
 RA Strausberg R.;
 RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 RN (7)
 RP SEQUENCE OF 66-431 FROM N.A.
 RX MEDLINE=84272706; PubMed=5896620;
 RA Verde P., Stoppelli M.P., Galetti P., di Nocera P., Blaai F.;
 RT "Identification and primary sequence of an unspliced human urokinase
 poly(A)+ RNA."
 RL Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).
 RN (8)
 RP SEQUENCE OF 21-177.
 RX MEDLINE=83055084; PubMed=6754569;
 RA Gunter W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E.,
 RA Floe L.;
 RT "The primary structure of high molecular mass urokinase from human
 urine. The complete amino acid sequence of the A chain."
 RL Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).
 RN (9)
 RP SEQUENCE OF 156-176 AND 179-224.
 RX MEDLINE=83003608; PubMed=6749491;
 RA Schaller J., Nick H., Rickl E.E., Gillesen D., Lergier W.,
 RA Studer R.O.;
 RT "Human low-molecular-weight urinary urokinase. Partial
 RT characterization and preliminary sequence data of the two polypeptide
 RT chains."
 RL Eur. J. Biochem. 125:251-257(1982).
 RN (10)
 RP SEQUENCE OF 158-410.
 RX MEDLINE=83055099; PubMed=6754572;
 RA Steffens G.J., Gunter W.A., Otting F., Frankus E., Floe L.;
 RT "The complete amino acid sequence of low molecular mass urokinase
 RT from human urine."
 RL Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).
 RN (11)
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE=96000858; PubMed=8591045;
 RA Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,
 RA Dobson C.M., Stuart D.I., Jones E.Y.;
 RT "The crystal structure of the catalytic domain of human
 RT urokinase-type plasminogen activator."
 RL Structure 3:681-691(1995).
 RN (12)
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.
 RX MEDLINE=20266327; PubMed=10805774;
 RA Sperl S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G.,
 RA Bode W., Magdalen V., Huber R., Moroder L.;
 RT "(4-aminomethyl)phenylguanidine derivatives as nonpeptidic highly
 RT selective inhibitors of human urokinase."
 RL Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).
 RN (13)
 RP STRUCTURE BY NMR.
 RX MEDLINE=89127526; PubMed=2536903;
 RA Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.;

RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-dimensional NMR.";
 RL Nature 337:579-582(1989).
 RN [14]
 RP STRUCTURE BY NMR OF 67-155.
 RX MEDLINE=93003110; PubMed=1327118;
 RA Li X., Smith R.A.G., Dobson C.M.;
 RT "Sequential 1H NMR assignments and secondary structure of the kringle domain from urokinase.";
 RL Biochemistry 31:9562-9571(1992).
 RN [15]
 RP STRUCTURE BY NMR OF 67-155.
 RX MEDLINE=94149701; PubMed=8107091;
 RA Li X., Bowman A.M., Lijonas M., Smith R.A.G., Dobson C.M.;
 RT "Solution structure of the kringle domain from urokinase-type plasminogen activator.";
 RL J. Mol. Biol. 235:1548-1559(1994).
 RN [16]
 RP VARIANT LEU-141.
 RX MEDLINE=96186279; PubMed=8652631;
 RA Yoshimoto M., Uehiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K., Sawasaki Y., Hamada K.;
 RT "Characterization of single chain urokinase-type plasminogen activator with a novel amino-acid substitution in the kringle structure.";
 RL Biochim. Biophys. Acta 1293:83-89(1996).
 RN [17]
 RP VARIANT LEU-141.
 RX MEDLINE=97218551; PubMed=9065988;
 RA Come B., Berczy M., Belin D.;
 RT "Detection of polymorphisms in the human urokinase-type plasminogen activator gene.";
 RL Thromb. Haemost. 77:434-435(1997).
 RN [18]
 RP ERRATUM.
 RA Come B., Berczy M., Belin D.;
 RL Thromb. Haemost. 78:973-973(1997).
 RN [19]
 RP VARIANT LEU-141.
 RX MEDLINE=97337920; PubMed=9194591;
 RA Turkmen B., Schmitt M., Schmalfeldt B., Trommler P., Hell W., Creutzburg S., Graeff H., Magdolen V.;
 RT "Molecular analysis of the genes encoding urokinase-type plasminogen activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer.";
 RL Electrophoresis 18:686-689(1997).
 CC -1- FUNCTION: POTENT PLASMINOGEN ACTIVATOR AND IS CLINICALLY USED FOR THERAPY OF THROMBOLYTIC DISORDERS.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.
 CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN.
 CC -1- PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used in Pulmonary Embolism (PE) to initiate fibrinolysis.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation- the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X02419; CA26268.1; -
 DR EMBL; M15476; AAA61253.1; -
 DR EMBL; D00244; BAA00175.1; -
 DR EMBL; D11433; BAA01919.1; -
 DR EMBL; X02760; CAA26535.1; -
 DR EMBL; AF377330; AAK53822.1; -

DR EMBL; BC013575; AAH13575.1; -
 DR EMBL; X03226; AAC97138.1; -
 DR EMBL; K02286; AAA61252.1; -
 DR EMBL; A21571; CAA01559.1; -
 DR EMBL; A18397; CAA01390.1; -
 DR PIR; A00931; UKRH.
 DR PIR; A32974; A32974.
 DR PDB; 1KDU; 31-OCT-93.
 DR PDB; 1LWV; 29-JAN-96.
 DR PDB; 1URK; 08-MAY-95.
 DR PDB; 1EJN; 17-MAY-00.
 DR MEROPS; S01.231; -
 DR GlycoSuiteDB; P00749; -
 DR Genew; HGNC:9052; PLAU.
 DR MIM; 191840; -
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00051; Kringle1; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRODOM; PD000395; Kringle1; 1.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.

Query Match 36.7%; Score 131; DB 1; Length 431;
 Best Local Similarity 40.3%; Pred. No. 6.9e-08;
 Matches 27; Conservative 7; Mismatches 25; Indels 8; Gaps 2;

QY 1 CFWDNCHLYREDQTSPPAGRLCLNWLDA-----OSGLASAPVSGAGNHSYCRNPDDSDP 53
 Db 70 CYEGNGHFFYRGKASITDWRGRLCPMNSATVLCQTYHAHNSDALQLGLGNGNCRNPD-NR 128
 QY 54 RGPWCYV 60
 Db 129 RRPWCYV 135

RESULT 11
 ID _UROK_PIG STANDARD; PRT; 442 AA.
 AC P04185;
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 GN PLAU.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxId=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=85087954; PubMed=6096832;
 RA Nagamine Y., Pearson D., Altus M.S., Reich E.;
 RT "cDNA and gene nucleotide sequence of porcine plasminogen activator.";
 RL Nucleic Acids Res. 12:9525-9541(1984).
 RN [2]
 RP REVISION TO 241.
 RA Nagamine Y.;
 RL Submitted (DEC-1986) to the PIR data bank.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X01648; CAA25806.1; -;
DR EMBL; X02724; CAA26511.1; -;
DR PIR; A00932; UNPG.
DR HSSP; P00749; IKDU.
DR MEROPS; S01.231; -;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001254; Ser_Protease_Try.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KM Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20
FT CHAIN 21 442 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 188 CHAIN A (BY SIMILARITY).
FT CHAIN 130 442 CHAIN B (BY SIMILARITY).
FT DOMAIN 29 65 EGF-LIKE.
FT DOMAIN 72 153 KRINGLE.
FT DOMAIN 154 189 CONNECTING PEPTIDE.
FT DOMAIN 442 442 SERINE PROTEASE.
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .).
FT DISULFID 33 41 BY SIMILARITY.
FT DISULFID 35 53 BY SIMILARITY.
FT DISULFID 55 64 BY SIMILARITY.
FT DISULFID 179 310 INTERCHAIN (BY SIMILARITY).
FT DISULFID 220 236 BY SIMILARITY.
FT DISULFID 228 299 BY SIMILARITY.
FT DISULFID 324 393 BY SIMILARITY.
FT DISULFID 356 372 BY SIMILARITY.
FT DISULFID 383 411 BY SIMILARITY.
FT ACT_SITE 235 235 CHARGE RELAY SYSTEM.
FT ACT_SITE 286 286 CHARGE RELAY SYSTEM.
FT ACT_SITE 387 387 CHARGE RELAY SYSTEM.
FT ACT_SITE 241 241 Q -> H (IN REF. 1; CAA25806).
FT CONFLICT 242 242 Q -> H (IN REF. 1; CAA26511).
FT CONFLICT 288 288 A -> G6 (IN REF. 1; CAA25806).
SQ SEQUENCE 442 AA; 49116 MW; 8322CFE501321EE CRC64;

Query Match 36.4%; Score 130; DB 1; Length 442;
Best local Similarity 44.8%; Pred. No. 9, 3e-08;
Matches 30; Conservative 4; Mismatches 25; Indels 8; Gaps 3;

QY 1 CFMDNGHLTREDDTSAPGLRLNWLDAQSGL-----ASAPVS---GAGNHSYCRANDEDP 53
DB 72 CFEGNHSHYRKANKNTVGRPLCPWNSATVLLNTYHARPDALQLGKGRNVCNRPD-NQ 130
QY 54 RGPWCYV 60
DB 131 RRPWCYV 137
RESULT 12
UROK_BOVIN
ID UROK_BOVIN STANDARD; PRT; 433 AA.
AC Q05589; Q8209;

DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UBA)
DE (U-plasminogen activator).
GN PLAU.
OC Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_Taxid=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Arteric endothelium;
RX MEDLINE=93216119; PubMed=835052;
RA Kretschmar U., Haendler B., Kojima S., Rifkin D.B.,
RA Schleming W.-D.;
RT "Bovine urokinase-type plasminogen activator and its receptor:
RT cloning and induction by retinoic acid";
RL Gene 125:177-183(1993).
RN [2]
RP SEQUENCE OF 12-433 FROM N.A.
RC Tissue=Kidney;
RA Ravn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators uPA
RT and tPA";
RL Int. Dairy J. 5:605-617(1995).
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-Val bond in
CC plasminogen to form plasmin.
CC -1- INDUCTION: BY RETINOIC ACID.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; L03546; AAA5149.1; -;
DR EMBL; X83801; CAA3796.1; -;
DR PIR; UN0560; UN0560.
DR HSSP; P00749; ILMW.
DR MEROPS; S01.231; -;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_Protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KM Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Zymogen.
FT SIGNAL 1 20
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 179 CHAIN A (BY SIMILARITY).
FT CHAIN 181 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 29 65 EGF-LIKE.
FT DOMAIN 72 153 KRINGLE.
FT DOMAIN 154 180 CONNECTING PEPTIDE.

```
FT DOMAIN 181 433 SERINE PROTEASE.
FT DISULFID 33 41 BY SIMILARITY.
FT DISULFID 35 53 BY SIMILARITY.
FT DISULFID 55 64 BY SIMILARITY.
FT DISULFID 170 301 INTERCHAIN (BY SIMILARITY).
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 290 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 277 277 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CONFLICT 189 189 A -> T (IN REF. 2).
SQ SEQUENCE 433 AA; 48730 MW; 4DE1B8D4DA47027A CRC64;

Query Match 35.6%; Score 127; DB 1; Length 433;
Best Local Similarity 41.2%; Pred. No. 2e-07; Mismatches 24; Indels 10; Gaps 3;
Matches 28; Conservative 6;

OY 1 CFWDNGHLYREDQTSPPAGLRCLNMLDAQSGL-----ASAPVSGAGNHSYCRNPDED 52
DB 72 CYQGGHSHYRKRARDLSGRPCIAM-DSPLYLLKMYHARSDAIQLGKHNVCNRPD-N 129
OY 53 RGPWCYV 60
DB 130 QRRWCYV 137

RESULT 13
UTRG DESRO STANDARD; PRT; 394 AA.
ID UTRG DESRO
AC P49150:
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator gamma precursor (EC 3.4.21.68) (DSPA gamma).
DE Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae; Desmodontinae; Desmodus.
OC NCBI_Taxid=9430;
OK [1]
RN RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=9203036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boisdol W., Bringmann P., Alagon A., Donner P., Schlenning W.D., "The plasminogen activator family from the salivary gland of the vampire bat Desmodus rotundus: cloning and expression.", Gene 105:229-237 (1991).
RL [2]
RN RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schlenning W.-D., Alagon A., Boisdol W., Bringmann P., Petri T., Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W., Donner P., "Plasminogen activators from the saliva of Desmodus rotundus (common vampire bat): unique fibrin specificity.", Ann. N.Y. Acad. Sci. 667:395-403 (1992).
CC -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC AGENT.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.
CC -!- SUBUNIT: MONOMER.
CC -!- SIMILARITY: BELONGS TO PERTINASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its
```

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M63990; AAA31595.1; -.
DR HSSP; P98119; 1A51.
DR MEROPS; S01.239; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000003; Kringle.
DR InterPro; IPR001254; Ser protease_Try.
DR Pfam; PF00051; Kringle_1.
DR Pfam; PF00089; Trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle_1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00200; TRYP_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
KW Kringle; Signal; Multigene family.
FT CHAIN 1 36 POTENTIAL.
FT SIGNAL 1 36
FT DOMAIN 37 394 SALIVARY PLASMINOGEN ACTIVATOR GAMMA.
FT ACT_SITE 142 394 KRINGLE.
FT ACT_SITE 189 189 SERINE PROTEASE.
FT ACT_SITE 238 238 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 345 345 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 45 126 BY SIMILARITY.
FT DISULFID 66 108 BY SIMILARITY.
FT DISULFID 97 121 BY SIMILARITY.
FT DISULFID 131 262 BY SIMILARITY.
FT DISULFID 174 190 BY SIMILARITY.
FT DISULFID 182 251 BY SIMILARITY.
FT DISULFID 276 351 BY SIMILARITY.
FT DISULFID 308 324 BY SIMILARITY.
FT DISULFID 341 369 BY SIMILARITY.
FT CARBOND 315 315 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 394 AA; 44105 MW; 9CDD6F52F3D81FCD CRC64;

Query Match 35.3%; Score 126; DB 1; Length 394;
Best Local Similarity 37.3%; Pred. No. 2.4e-07; Mismatches 26; Indels 8; Gaps 2;
Matches 25; Conservative 8;

OY 1 CFWDNGHLYREDQTSPPAGLRCLNML-----LDAQSGLASAPVSGAGNHSYCRNPDED 53
DB 45 CYKDDQVTRGTWSTSSGAGCINMNSNLIIRTYNGRMPEAVKLGUNHNVCNRPDGS 104
OY 54 RGPWCYV 60
DB 105 K-PWCYV 110

RESULT 14
ID HGFL MOUSE STANDARD; PRT; 716 AA.
AC P26928;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor-like protein precursor (Macrophage stimulatory protein) (MSP).
DE scumulatory protein (MSP).
GN MST1 OR HGFL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RN RP SEQUENCE FROM N.A.
```

```

BC STRAIN=BAUB/C; TISSUE=Liver;
RA MEDLINE=92002017; PubMed=1832957;
RT Friezen Degen S.J., Stuart L.A., Han S., Jamison C.S.;
RT "Characterization of the mouse cDNA and gene coding for a hepatocyte
RT growth factor-like protein: expression during development.";
RL Biochemistry 30:9781-9791(1991).
CC -1- FUNCTION: PROBABLY HAS NO PROTEOLYTIC ACTIVITY, SINCE CRUCIAL AA
CC CHARACTERISTIC OF SERINE PROTEASES CATALYTIC SITES ARE NOT
CC CONSERVED.
CC -1- TISSUE SPECIFICITY: LIVER. LOWER LEVELS IN LUNG, PLACENTA AND
CC ADRENAL.
CC -1- DEVELOPMENTAL STAGE: IS EXPRESSED AT LOW LEVELS DURING GESTATION.
CC JUST BEFORE BIRTH THE LEVEL INCREASES DRAMATICALLY AND REMAINS
CC STABLE AFTERWARDS.
CC -1- PTM: MAY BE CLEAVED AFTER AA 488, OR CLEAVED INTO TWO SEPARATE
CC HELD TOGETHER BY DISULFIDE BONDS, OR CLEAVED INTO TWO SEPARATE
CC POLYPEPTIDES.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M74180; AAA50166.1; -.
DR EMBL; M74181; AAA50167.1; -.
DR HSSP; P00747; IKRN.
DR MEROPS; S01.975; -.
DR MGD; MGI:96080; Hgf1.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kring1.
DR InterPro; IPR003609; Pan_ap.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF000024; PAN; 1.
DR Pfam; PF000051; kring1; 4.
DR Pfam; PF000089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; Kring1.
DR ProDom; PD000395; Kring1; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS02440; TRYPSIN_DOM; 1.
KW Kring1; Glycoprotein; Serine protease homolog; Repeat; Signal.
FT SIGNAL 1 18
FT CHAIN 19 716
FT DOMAIN 19 109
FT DOMAIN 110 186
FT DOMAIN 191 268
FT DOMAIN 292 370
FT DOMAIN 379 457
FT DOMAIN 489 716
FT DISULFID 56 78
FT DISULFID 60 66
FT DISULFID 110 186
FT DISULFID 131 169
FT DISULFID 157 181
FT DISULFID 191 268
FT DISULFID 194 333
FT DISULFID 212 251
FT DISULFID 240 263
FT DISULFID 292 370
FT DISULFID 313 352
FT DISULFID 341 364
FT DISULFID 379 457
FT DISULFID 400 440
BY SIMILARITY.

```

```

FT DISULFID 428 452
FT DISULFID 477 593
FT DISULFID 512 528
FT DISULFID 607 672
FT DISULFID 637 651
FT DISULFID 662 690
FT CARBOHYD 72 72
FT CARBOHYD 173 173
FT CARBOHYD 305 305
FT CARBOHYD 620 620
FT CONFLICT 19 19
SQ SEQUENCE 716 AA; 80588 MW; BBCE02EF85213ACC CRC64;
Query Match 35.3%; Score 126; DB 1; Length 716;
Best Local Similarity 41.9%; Pred. No. 4,4e-07;
Matches 26; Conservative 3; Mismatches 29; Indels 4; Gaps 2;
QY 1 CFWDNGHLYREDQTSPPGRLCLNW--LDAQSGLASAPVSGAGNHSYCRNPDEDPGFW 57
DB 110 CIMDNGSYRTGTVARTAGLGPQAWSRFPNDHKYTPPKNGL-EEFNCRNPDGPGFW 168
QY 58 CY 59
DB 169 CY 170
RESULT 15
PLMN RAT STANDARD; PRT; 169 AA.
AC 001177;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Plasminogen (EC 3.4.21.7) (Fragment).
GN Plg.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA MEDLINE=91250378; PubMed=1645711;
RX Kanalas J.J., Makker S.P.;
RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a
RT receptor site for plasminogen.";
RL J. Biol. Chem. 266:10825-10829(1991).
CC -1- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC AND INFLAMMATION; IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC ACTIVATOR, COLLAGENASES AND SEVERAL CYTOKINES, SUCH
CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC LAMININ AND VON WILLEBRAND FACTOR.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -1- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC -1- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```



```
CC -----
DR EMBL; M62832; AAA41884.1; -.
DR PIR; A40522; A40522.
DR HSSP; P00747; 1PMK.
DR MEROPS; S01.233; -.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; kringle; 2.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS0240; TRYPsin_DOM; PARTIAL.
DR PROSITE; PS00134; TRYPsin_HIS; PARTIAL.
DR PROSITE; PS00135; TRYPsin_SER; PARTIAL.
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Repeat.
FT NON_TER 1
FT DOMAIN 1
FT DOMAIN <1 10 KRINGLE 3 (BY SIMILARITY).
FT DOMAIN 34 112 KRINGLE 4 (BY SIMILARITY).
FT DOMAIN 139 >169 KRINGLE 5 (BY SIMILARITY).
FT DISULFID 34 112 BY SIMILARITY.
FT DISULFID 55 95 BY SIMILARITY.
FT DISULFID 83 107 BY SIMILARITY.
FT NON_TER 169 169
SQ SEQUENCE 169 AA; 18401 MW; 77A54214C9D010C CRC64;

Query Match 34.5%; Score 123; DB 1; Length 169;
Best Local Similarity 37.5%; Pred. No. 2,2e-07;
Matches 24; Conservative 9; Mismatches 25; Indels 6; Gaps 2;

QY 1 CFNDNGHYREDQSPAPGRCIANLDA-----QSGLASAPVSGAGNHSYCRNPDPDPG 55
DB 34 CIGQNGKSYKGTSSITWTGKCKQSWMTSPHSHSKTPANFPDSGL-EMNTCRNPNDORG 92
QY 56 PWCY 59
DB 93 PWCF 96
```

Search completed: April 7, 2003, 09:16:57
Job time : 3.74428 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:25 ; Search time 16.0915 Seconds
(without alignments)
768.284 Million cell updates/sec

Title: US-10-057-951-2_COPY_25_84

Perfect score: 357
Sequence: 1 CFWDNGHLYREDQTSPPAPGL.....GNHSYCRNPDEDPGPGWCYV 60

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	357	100.0	263	4	000318
2	357	100.0	263	4	096FE7
3	155	43.4	562	6	08SQ23
4	146	40.9	516	4	09BU99
5	146	40.9	653	11	08VC84
6	136	38.1	559	11	091VP2
7	131	36.7	154	4	096S88
8	130.5	36.6	717	13	P70006
9	128	35.9	716	11	F70521
10	127	35.6	157	6	09TVAR
11	127	35.6	313	13	09PU78
12	127	35.6	395	4	09BZM1
13	126	35.3	716	11	091XG8
14	122.5	34.3	716	13	0916R1
15	122	34.2	704	13	090865
16	121	33.9	810	4	Q15146

17	120	33.6	812	11	09R0M3	09R0M3
18	118	33.1	728	6	09BH09	09BH09
19	118	33.1	806	6	018783	018783
20	117.5	32.9	399	4	096GL8	096GL8
21	117.5	32.9	420	4	09BTP9	09BTP9
22	113	31.7	452	13	090Y90	090Y90
23	112.5	31.5	334	6	046507	046507
24	111	31.1	385	5	025101	025101
25	109	30.5	202	13	090675	090675
26	109	30.5	567	4	013208	013208
27	109	30.5	685	5	024488	024488
28	108.5	30.4	615	6	097507	097507
29	108.5	30.4	812	11	091M05	091M05
30	108	30.3	115	13	042341	042341
31	107.5	30.1	1145	5	09BKL8	09BKL8
32	106.5	29.8	420	13	090504	090504
33	106	29.7	359	6	08WMR1	08WMR1
34	106	29.7	378	13	090WP0	090WP0
35	105	29.4	710	13	091402	091402
36	105	29.4	726	13	090978	090978
37	104	29.1	560	4	014520	014520
38	103	28.9	608	13	09PTW7	09PTW7
39	102.5	28.7	597	11	035727	035727
40	102	28.6	385	13	090WS2	090WS2
41	95.5	26.8	113	4	09UTR5	09UTR5
42	94	26.3	709	13	090ZN6	090ZN6
43	93.5	26.2	594	5	P91823	P91823
44	93	26.1	214	6	09XT70	09XT70
45	92.5	25.9	648	4	Q9H1V4	Q9H1V4

ALIGNMENTS

RESULT 1

AC 000318 PRELIMINARY; PRT; 263 AA.
ID 000318
AC 000318;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE WUSGC:DJ515N1.2 protein.
GN WUSGC:DJ515N1.2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Du Z., Scheet P., Harper M., PAC clone RP3-515N1."
RT "The sequence of H. sapiens clone RP3-515N1."
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Waterston R.
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC002073; AAB54054.1; -
DR HSSP; P00749; IKDU.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE; FALSE_NEG.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SMO0130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
SQ SEQUENCE 263 AA; 28248 MW; 197C3EEBE54A242 CRC64;

Query Match 100.0%; Score 357; DB 4; Length 263;
Best Local Similarity 100.0%; Pred. No. 2,4e-34;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CFWDNGHLYREDQTSPPAPGLRLNWLDAQSGIASAPVSGAGNHSYCRNPDEDPGPGWCYV 60
|||||

```
DB 25 CFWDNGHLVREDQTSAPGLRCLNMLDAQSGLASAPVSGAGNHSYCRNPDEDPGPGWCYV 84
RESULT 2
ID 096FE7 PRELIMINARY; PRT; 263 AA.
AC 096FE7;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Unknown (protein for MGCL17330).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; BC011049; AAH11049.1; -.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR ProDom; PD000395; Kringle; 1.
DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_1.
DR PROSITE; PS00070; KRINGLE_2; 1.
SQ SEQUENCE 263 AA; 28234 MW; 197C3EBE88F242 CRC64;

Query Match 100.0%; Score 357; DB 4; Length 263;
Best Local Similarity 100.0%; Pred. No. 2,4e-34;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFWDNGHLVREDQTSAPGLRCLNMLDAQSGLASAPVSGAGNHSYCRNPDEDPGPGWCYV 60
DB 25 CFWDNGHLVREDQTSAPGLRCLNMLDAQSGLASAPVSGAGNHSYCRNPDEDPGPGWCYV 84
RESULT 3
ID 08SQ23 PRELIMINARY; PRT; 562 AA.
AC 08SQ23;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE T-plasminogen activator.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_Taxid=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ENMEL ORGAN;
RA Ding Y., Xue J., Bartlett J.D.;
RT "T-plasminogen activator in tooth tissue."
RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF364605; AAM00297.1; -.
SQ SEQUENCE 562 AA; 63668 MW; F9E6BAC77CB101B8 CRC64;

Query Match 43.4%; Score 155; DB 6; Length 562;
Best Local Similarity 43.5%; Pred. No. 4,1e-10;
Matches 30; Conservative 9; Mismatches 18; Indels 12; Gaps 3;

QY 1 CFWDNGHLVREDQTSAPGLRCLNMLDAQSGLASAPVSGAGNHSYCRNPDEDPGPGWCYV 51
DB 127 CFWDNGHLVREDQTSAPGLRCLNMLDAQSGLASAPVSGAGNHSYCRNPDEDPGPGWCYV 184
QY 52 DPGPGWCYV 60
DB 185 DSK-PWCYI 192
RESULT 4
Q9BU99
```

```
ID Q9BU99 PRELIMINARY; PRT; 516 AA.
AC Q9BU99;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Similar to plasminogen activator, tissue.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SKIN;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPsin FAMILY.
CC EMBL; BC002795; AAH02795.1; -.
DR HSSP; P00750; IASH.
DR MEROPS; S01.232; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00051; Kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF_1ke; 1.
DR SMART; SM00001; EGF_1ke; 1.
DR SMART; SM00130; KR_2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Serine protease.
SQ SEQUENCE 516 AA; 57370 MW; BAB31901FDC96800 CRC64;

Query Match 40.9%; Score 146; DB 4; Length 516;
Best Local Similarity 43.5%; Pred. No. 4,4e-09;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

QY 1 CFWDNGHLVREDQTSAPGLRCLNMLDAQSGLASAPVSGAGNHSYCRNPDEDPGPGWCYV 51
DB 81 CFWDNGHLVREDQTSAPGLRCLNMLDAQSGLASAPVSGAGNHSYCRNPDEDPGPGWCYV 138
QY 52 DPGPGWCYV 60
DB 139 DSK-PWCYV 146
RESULT 5
Q8VCS4 PRELIMINARY; PRT; 653 AA.
ID Q8VCS4;
AC Q8VCS4;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 70.6 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Strausberg R.;
```

RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC019376; AA019376.1; -
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000742; EGF_2.
 DR InterPro: IPR001438; EGF_1.
 DR InterPro: IPR000883; Fibrinctn.
 DR InterPro: IPR000562; FN Type_1.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR01254; Ser_protease_Try.
 DR Pfam: PF00008; EGF_2.
 DR Pfam: PF00039; fn1; 1.
 DR Pfam: PF00040; fn2; 1.
 DR Pfam: PF00051; kringle; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00010; EGFLOOD.
 DR PRINTS: PR00013; FNTYPE1.
 DR PRINTS: PR00018; KRINGLE.
 DR ProDom: PD000395; Kringle_1.
 DR ProDom: PD000995; FN_Type_1; 1.
 DR SMART: SM00181; EGF_2.
 DR SMART: SM00058; FN1; 1.
 DR SMART: SM00059; FN2; 1.
 DR SMART: SM00130; KR; 1.
 DR SMART: SM00020; TRYP_Spc; 1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
 DR PROSITE: PS01253; FIBRONECTIN_1; UNKNOWN_1.
 DR PROSITE: PS00023; FIBRONECTIN_2; UNKNOWN_1.
 DR PROSITE: PS00021; KRINGLE_1; UNKNOWN_1.
 DR PROSITE: PS50070; KRINGLE_2; 1.
 DR PROSITE: PS50240; TRYPsin_DOM; 1.
 DR PROSITE: PS00134; TRYPsin_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPsin_SER; UNKNOWN_1.
 KW Hypothetical protein.
 SQ SEQUENCE 653 AA; 70553 MW; FE18D90174ED6FDD CRC64;

Query Match 40.9%; Score 146; DB 11; Length 653;
 Best Local Similarity 47.1%; Pred. No. 5.6e-09;
 Matches 32; Conservative 7; Mismatches 19; Indels 10; Gaps 3;
 QY 1 CPWDNGHLYREQOTSPAPGLRCLNM-----LDAQSLASAPVSGAGNHSYGRNDE 52
 DB 283 CFLGGETRGVASTAASGLSCLANNSDLVQELHVDV-VAAAVLLGLGPHAYCENPDK 341
 QY 53 PRGPMCV 60
 DB 342 ER-PWCYV 348

RESULT 6
 Q91VP2 PRELIMINARY; PRT; 559 AA.
 AC Q91VP2; 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Similar to plasminogen activator, tissue.
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BREAST TUMOR;
 RA Strusberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC011256; AA011256.1; -
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000083; Fibrinctn.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR001254; Set_protease_Try.

DR Pfam: PF00008; EGF_1.
 DR Pfam: PF00039; fn1; 1.
 DR Pfam: PF00051; kringle; 2.
 DR Pfam: PF00089; trypsin; 1.
 DR ProDom: PD000395; Kringle_2.
 DR SMART: SM00181; EGF_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
 DR PROSITE: PS01253; FIBRONECTIN_1; UNKNOWN_1.
 DR PROSITE: PS00021; KRINGLE_1; UNKNOWN_2.
 DR PROSITE: PS50070; KRINGLE_2; 2.
 DR PROSITE: PS50240; TRYPsin_DOM; 1.
 DR PROSITE: PS00134; TRYPsin_HIS; UNKNOWN_1.
 DR PROSITE: PS00135; TRYPsin_SER; UNKNOWN_1.
 KW Hydrolyase, Serine protease
 SQ SEQUENCE 559 AA; 63122 MW; 8CCEB2BD94514D9 CRC64;

Query Match 38.1%; Score 136; DB 11; Length 559;
 Best Local Similarity 40.6%; Pred. No. 7.3e-08;
 Matches 28; Conservative 8; Mismatches 21; Indels 12; Gaps 3;
 QY 1 CPWDNGHLYREQOTSPAPGLRCLNM-----LDAQSLASAPVSGAGNHSYGRNDE 51
 DB 124 CFEEOGITYRGTWSTAESGACINM--NSSVLSLKPYARRRNPAIKLGLGNHYCRNPD 181
 QY 52 DPRGPMCV 60
 DB 182 DLK-PWCYV 189

RESULT 7
 Q96SE8 PRELIMINARY; PRT; 154 AA.
 AC Q96SE8; 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Urokinase-type plasminogen activator amino-terminal
 DE fragment.
 GN ANF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Fu J., Bai X., Ruan C.;
 RT "Cloning and expression of the amino-terminal fragment of human
 RT urokinase-type plasminogen activator."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Bai X., Wang W., Xi X., Ruan C.;
 RT "Overexpression of the amino-terminal fragment of human urokinase-type
 RT plasminogen activator in breast cancer cells results in decreased
 RT tumor invasion, growth and angiogenesis."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY029537; AA029537.1; -
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000001; Kringle.
 DR Pfam: PF00051; kringle; 1.
 DR ProDom: PD000395; Kringle_1.
 DR SMART: SM00181; EGF_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE: PS00021; KRINGLE_1; UNKNOWN_1.
 DR PROSITE: PS50070; KRINGLE_2; 1.
 KW Kinase.
 SQ SEQUENCE 154 AA; 17305 MW; A3CCF2FCFF505572 CRC64;

Query Match 36.7%; Score 131; DB 4; Length 154;
 Best Local Similarity 40.3%; Pred. No. 7.7e-08;
 Matches 27; Conservative 7; Mismatches 25; Indels 8; Gaps 2;

OY 1 CFWDNGHLYREDQTSAPAGLACLNWLDA-----QSGLASAPVSGAGNHSYCRNPDED 53
 DB 70 CYEGNGHFRGASTDTPGRCPLPWNASATVLOQTYHARSBLQGLGKKNYCRNPD-NR 128
 OY 54 RGPWCYV 60
 DB 129 RRPWCYV 135

RESULT 8

P70006 PRELIMINARY; PRT; 717 AA.
 AC P70006;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hepatocyte growth factor-like protein precursor.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
 OC Xenopodinae; Xenopus.
 NC NCB1_TaxID=8355;

RN [1]
 RC SEQUENCE FROM N.A.
 RP TISSUE=KIDNEY;
 RX MEDLINE=96404125; PubMed=8808403;
 RA Abergner F., Schmidt G., Richter K.;
 RT "The Xenopus homologue of hepatocyte growth factor-like protein is
 RT specifically expressed in the presumptive neural plate during
 RT gastrulation.";
 RL Mech. Dev. 54:23-37(1996).
 DR EMBL; Y08734; CAA69989.1; -.
 DR HSSP; P00747; ICEA.
 DR MEROPS; S01.977; -.

DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR003014; PAN.
 DR InterPro; IPR003609; Pan_app.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00051; Kringle; 4.
 DR Pfam; PF00024; PAN; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00018; KRINGLE.
 DR PRODOM; PD000395; Kringle; 4.
 DR SMART; SM00130; KR; 4.
 DR SMART; SM00473; PAN_AP; 1.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS00021; KRINGLE_1; 3.
 DR PROSITE; PS50070; KRINGLE_2; 4.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 KW Hydrolyase; Serine protease; Signal.
 FT SIGNAL 1
 FT CHAIN 29 717 HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN.
 SQ SEQUENCE 717 AA; 82017 MW; 6F877A432C8DD54 CRC64;

Query Match 36.6%; Score 130.5; DB 13; Length 717;
 Best Local Similarity 38.7%; Pred. No. 4.2e-07;
 Matches 24; Conservative 7; Mismatches 28; Indels 3; Gaps 1;

OY 1 CFWDNGHLYREDQTSAPAGLACLNWLDAQSGT---ASAPVSGAGNHSYCRNPDEDPRGPW 57
 DB 378 CYHNGELIYRGVSKTRKGIWCRWEKQNDLELSLAQPIVPLENYCRNPDRSHGPW 437
 OY 58 CY 59
 DB 438 CY 439

RESULT 9

P70521 PRELIMINARY; PRT; 716 AA.
 AC P70521;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Macrophage stimulating protein precursor.
 OS Rattus norvegicus (Rat).
 NC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NC NCB1_TaxID=10116;

RN [1]
 RC SEQUENCE FROM N.A.
 RP TISSUE=LIVER;
 RX MEDLINE=97011126; PubMed=8858136;
 RA Ohshtiro K., Iwama A., Matsuno K., Ezaki T., Sakamoto O., Hamaguchi I.,
 RA Takasu N., Suda T.;
 RT "Molecular cloning of Rat Macrophage-stimulating protein and its
 RT involvement in the Male Reproductive System."
 RL Biochem. Biophys. Res. Commun. 227:273-280(1996).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.

DR EMBL; X95096; CAA64473.1; -.
 DR HSSP; P00747; ICRN.
 DR MEROPS; S01.975; -.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR003014; PAN.
 DR InterPro; IPR003609; Pan_app.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00051; Kringle; 4.
 DR Pfam; PF00024; PAN; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR PRODOM; PD000395; Kringle; 4.
 DR SMART; SM00130; KR; 4.
 DR SMART; SM00473; PAN_AP; 1.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS00021; KRINGLE_1; 4.
 DR PROSITE; PS50070; KRINGLE_2; 4.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 KW Hydrolyase; Serine protease; Signal.
 FT SIGNAL 1
 FT CHAIN 31
 SQ SEQUENCE 716 AA; 80733 MW; 06B7DF3EF56D922F CRC64;

Query Match 35.9%; Score 128; DB 11; Length 716;
 Best Local Similarity 41.9%; Pred. No. 8.3e-07;
 Matches 26; Conservative 3; Mismatches 29; Indels 4; Gaps 2;

OY 1 CFWDNGHLYREDQTSAPAGLACLNW---LDAQSGLASAPVSGAGNHSYCRNPDEDPRGPW 57
 DB 110 CIMDNGASYRGTVARTADGLPCQASRRFPNDHKYTPRPKGL-BENFCRNPDPGRGPW 168
 OY 58 CY 59
 DB 169 CY 170

RESULT 10
 O9TVAB PRELIMINARY; PRT; 157 AA.

AC O9TVAB;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Urokinase plasminogen activator (Fragment).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 OC Bovidae; Bovinae; Bos.
 NC NCB1_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SKLETAL MUSCLE;
 RX MEDLINE=21071388; PubMed=11204721;
 RA Balcerzak D., Querengueser L., Dixon W.T., Baracos V.E.;

"Coordinate expression of matrix-degrading proteinases and their RT activators and inhibitors in bovine skeletal muscle.";
RT J. Anim. Sci. 79:94-107(2001).
DR EMBL: AF144761; AAD30301.1; -.
DR HSSP: P00749; IURK.
DR MEROPS; S01.231; -.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle_1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
KW Kinase.
FT NON_TER 1 157 157
SQ SEQUENCE 157 AA; 17858 MW; A768D6C72C1FBFB7 CRC64;

Query Match 35.6%; Score 127; DB 6; Length 157;
Best Local Similarity 41.2%; Pred. No. 2.3e-07;
Matches 28; Conservative 6; Mismatches 24; Indels 10; Gaps 3;

QY 1 CFWDNGHLYREDQTSPPAGRLCLMWLDAQSGL-----ASAPVSGAGNHSYCRNPDE 52
DB 38 CYGNGSHSYRKARNDLSGRPCLAW-DSPYTLGMVYHAHSDAQLGLGKHNVCNPDD-N 95
QY 53 PRGWCYV 60
DB 96 QRRPCYV 103

RESULT 11
Q9PU78 PRELIMINARY; PRT; 313 AA.
ID Q9PU78
AC Q9PU78;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hepatocyte growth factor-like protein (Fragment).
OS Crocodylus niloticus (Nilie crocodile) (African crocodile).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodyliidae; Crocodylinae; Crocodylus.
OX NCBI_Taxid=8501;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=20022983; PubMed=10555283;
RA Hughes S., Zelus D., Mouchroud D.;
RT "Warm-blooded isochore structure in nilie crocodile and turtle.";
RL Mol. Biol. Evol. 16:1521-1527(1999).
DR EMBL; AJ011396; CAB56422.1; -.
DR HSSP; P00747; IHPJ.
DR MEROPS; S01.977; -.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_Protease_Try.
DR Pfam; PF00051; kringle_2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
KW Hydrolyase; Serine protease.
FT NON_TER 1 157 157
SQ SEQUENCE 313 AA; 34793 MW; 8E084704958B5A2 CRC64;

Query Match 35.6%; Score 127; DB 13; Length 313;
Best Local Similarity 36.5%; Pred. No. 4.7e-07;

Matches 23; Conservative 8; Mismatches 28; Indels 4; Gaps 2;

QY 1 CFWDNGHLYREDQTSPPAGRLCLMWLDAQSGLAS-APVSGAGN-----SYCRNPDEPRGP 56
DB 30 CYGNGSHLYRGHTSKTRRGVTCQKWSQSPHYVPSPTTHRAAHLDENVGKPNDSNGP 89
QY 57 WCY 59
DB 90 WCY 92

RESULT 12
Q9BZWI PRELIMINARY; PRT; 395 AA.
ID Q9BZWI
AC Q9BZWI;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Neonatal thrombolytic agent alpha-form (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Dou D.;
RT "A brain-type plasminogen activator.";
RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.
DR EMBL; AF260825; AAK11956.1; -.
DR HSSP; P00750; IPR2.
DR MEROPS; S01.232; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000083; Fibrinctn.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_DOM; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease.
FT NON_TER 395 395
SQ SEQUENCE 395 AA; 44323 MW; 3FBD4A2F0B7C11C8 CRC64;

Query Match 35.6%; Score 127; DB 4; Length 395;
Best Local Similarity 39.1%; Pred. No. 5.9e-07;
Matches 27; Conservative 7; Mismatches 23; Indels 12; Gaps 3;

QY 1 CFWDNGHLYREDQTSPPAGRLCLMW-----LDAQSGLASAPVSGAGNHSYCRNPDE 51
DB 48 CYGNGSHLYRGHTSHLSGASCLPNNSMILLGKYTAQN--PSAALGLGKHNVCNPDDG 105
QY 52 DPRGWCYV 60
DB 106 DAK-PWCHV 113

RESULT 13
Q91XG8 PRELIMINARY; PRT; 716 AA.
ID Q91XG8
AC Q91XG8;

DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hepatocyte growth factor-like.
GN HGF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Struhsberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010551; AAH10551.1; -.
DR MGD; MGI:96080; Hgf1.
DR InterPro; IPR00001; Kring1e.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; Kring1e; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR ProDom; PD00395; Kring1e; 4.
DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_4.
DR PROSITE; PS0070; KRINGLE_2; 4.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
KW Hydrolyase; Serine protease.
SQ SEQUENCE 716 AA; 80693 MW; 12474C48A7D4B46D CRC64;

Query Match 35.3%; Score 126; DB 11; Length 716;
Best Local Similarity 41.9%; Pred. No. 1.4e-06;
Matches 26; Conservative 3; Mismatches 29; Indels 4; Gaps 2;
QY 1 CFWDNGLYREDQTSAPGLRCLNWLDAQSGLASAPVSGAGHSYCRNPDEPRGPM 57
DB 110 CIMGNGVSYRGTVARTAGGLPCQAWSRFPNDHKYTPPKXGL-BENFCRNPDDGPRGPM 168
QY 58 CY 59
DB 169 CY 170

RESULT 14
Q91691 PRELIMINARY; PRT; 716 AA.
AC Q91691;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Growth factor liverline.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxId=8335;
RN [1]
RP SEQUENCE FROM N.A.
RA Ruiz i Altaba A.; They C.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U57455; AAB52574.1; -.
DR HSSP; P00747; ICEA.
DR InterPro; IPR00001; Kring1e.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; Kring1e; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PROSITE; PS00018; KRINGLE.
DR ProDom; PD000395; Kring1e; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; Tryp_Spc; 1.

DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
KW Hydrolyase; Serine protease.
SQ SEQUENCE 716 AA; 81971 MW; 508376A0E4398798 CRC64;

Query Match 34.3%; Score 122.5; DB 13; Length 716;
Best Local Similarity 35.5%; Pred. No. 3.7e-06;
Matches 22; Conservative 8; Mismatches 29; Indels 3; Gaps 1;
QY 1 CFWDNGLYREDQTSAPGLRCLNWLDAQSGLASAPVSGAGHSYCRNPDEPRGPM 57
DB 377 CYHNGELYSGRVSKTRKIKCRWEKRNDELSDQYLVPLBENYCRNPDRSHGPM 436
QY 58 CY 59
DB 437 CY 438

RESULT 15
Q90865 PRELIMINARY; PRT; 704 AA.
AC Q90865;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Hepatocyte growth factor-like/macrophage stimulating protein.
GN HGF/MSP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxId=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=96029010; PubMed=7554499;
RA They C., Sharpe M.J., Batley S.J., Stern C.D., Gherardi E.;
RT "Expression of HGF/SF, HGF/MSP and c-met suggests new functions
RT during early chick development.";
RC Dev. Genet. 17:90-101(1995).
CL -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC EMBL; X84043; CAA5862.1; -.
DR HSSP; P00747; ICEA.
DR MEROPS; S01.977; -.

DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kring1e.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; Kring1e; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kring1e; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS0070; KRINGLE_2; 4.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
KW Hydrolyase; Serine protease.
SQ SEQUENCE 704 AA; 79341 MW; CAB0D8CC41367C37 CRC64;

Query Match 34.2%; Score 122; DB 13; Length 704;
Best Local Similarity 37.3%; Pred. No. 4.2e-06;
Matches 25; Conservative 6; Mismatches 22; Indels 14; Gaps 2;
QY 1 CFWDNGLYREDQTSAPGLRCLNWLDAQSGLASAPVSGAGHSYCRNPDEPRGPM 52
DB 108 CIVANGTSYRGTRDTERGLRCQHW-----CATPHHRFLPGLRNGLEENYCRNPDRD 161

Mon Apr 7 10:22:32 2003

us-10-057-951-2_copy_25_84.rpt

Page 7

OY 53 PRGPNCT 59
| | | | |
DB 152 KRGPACT 158

Search completed: April 7, 2003, 09:21:36
JOB time : 17.0915 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:24 ; Search time 11.2266 Seconds
(without alignments)
712.151 Million cell updates/sec

Title: US-10-057-951-2_COPY_25_84
Perfect score: 357
Sequence: 1 CFMDNGHLYREDQTSPPAGL.....GNHSYCNDEPDPRGWCYV 60

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID52/gcgdata/geneseq/geneeqp-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneeq/geneeqp-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneeq/geneeqp-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneeq/geneeqp-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneeq/geneeqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	357	100.0	263	20 AAY05219	Kringlet protein s
2	357	100.0	263	20 AAM87769	Human tissue plas
3	357	100.0	263	22 AAM93748	Human polypeptide,
4	357	100.0	263	22 AAE00300	Human tissue-plasm
5	357	100.0	263	23 AAB06149	Human PRO264 polyp
6	351	98.3	263	21 AAB43237	Human ORFX ORF3001
7	343	96.1	286	20 AAY05220	Kringlet protein s
8	186	52.1	56	20 AAY12615	Human 5' EST seque
9	185	51.8	55	20 AAY12397	Human 5' EST seque
10	183	51.3	39	19 AAM72641	Nervous glia cell

11	183	51.3	39	19 AAM72640	Nervous glia cell
12	149	41.7	527	19 AAM54154	t-PA mutant (N142S
13	149	41.7	527	19 AAM54157	t-PA mutant (N142S
14	148	41.5	390	9 AAP82587	Modified tissue pl
15	148	41.5	483	16 AAR70889	Human tissue PA va
16	148	41.5	483	16 AAR70877	Human tissue PA va
17	148	41.5	483	16 AAR70878	Human tissue PA va
18	148	41.5	483	16 AAR70879	Human tissue PA va
19	148	41.5	483	16 AAR70880	Human tissue PA va
20	148	41.5	483	16 AAR70881	Human tissue PA va
21	148	41.5	483	16 AAR70882	Human tissue PA va
22	148	41.5	483	16 AAR70883	Human tissue PA va
23	148	41.5	483	16 AAR70884	Human tissue PA va
24	148	41.5	483	16 AAR70885	Human tissue PA va
25	148	41.5	483	16 AAR70886	Human tissue PA va
26	148	41.5	483	16 AAR70887	Human tissue PA va
27	148	41.5	483	16 AAR70888	Human tissue PA va
28	148	41.5	483	16 AAR70890	Human tissue PA va
29	148	41.5	527	13 AAR21598	Human tPA variant
30	148	41.5	527	14 AAR44812	Human tPA variant
31	148	41.5	527	14 AAR44816	Human tPA variant
32	148	41.5	527	14 AAR44817	Human tPA variant
33	148	41.5	527	16 AAR70874	Human t-PA variant
34	148	41.5	527	16 AAR70875	Human t-PA variant
35	148	41.5	527	16 AAR70876	Human t-PA variant
36	148	41.5	527	16 AAR70898	Human t-PA variant
37	148	41.5	527	16 AAR70899	Human t-PA variant
38	148	41.5	527	16 AAR70900	Human t-PA variant
39	148	41.5	527	16 AAR70901	Human t-PA variant
40	148	41.5	527	16 AAR70902	Human t-PA variant
41	148	41.5	527	16 AAR70903	Human t-PA variant
42	148	41.5	527	16 AAR70904	Human t-PA variant
43	148	41.5	527	16 AAR70905	Human t-PA variant
44	148	41.5	527	16 AAR70906	Human t-PA variant
45	148	41.5	527	16 AAR70907	Human t-PA variant

ALIGNMENTS

RESULT 1	AA105219	
ID	AA105219	standard, Protein; 263 AA.
XX		
AC	AA105219;	
XX		
DT	17-JUN-1999	(first entry)
XX		
DE	Kringlet protein sequence.	
XX		
KM	Kringlet; cancer; inflammation; autoimmunity; rheumatoid arthritis; AIDS;	
KM	CNS inflammation; cerebellar degeneration; Alzheimer's disease; ashma;	
KM	Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;	
KM	neurological abnormality; ischemia reperfusion injury; ischaemic injury;	
KM	cardiovascular disease; kidney disease; liver disease; aplastic anaemia;	
KM	myocardial infarction; hypotension; hypertension; allergy; infection;	
KM	myelodysplastic syndrome; haematologic abnormality; diagnosis; therapy;	
KM	male pattern baldness.	
XX		
OS	Homo sapiens.	
XX		
PN	WO911788-A1.	
XX		
PD	11-MAR-1999.	
XX		
PF	02-SEP-1998;	98MO-US18270.
XX		
PR	01-SEP-1998;	98US-0144889.
XX		
PR	02-SEP-1997;	97US-0056032.
XX		
PA	(SMIK) SMITHKLINE BEECHAM CORP.	
XX		
PI	Albone EF, Kikly KK;	

```

XX MPI: 1999-214707/18.
DR N-PSDB; AAX28354.
XX
XX New kringle1 polypeptides and polynucleotides
XX
XX Claim 1; Page 31-32; 42pp; English.
XX
CC This sequence is a Kringle1 polypeptide of the invention.
CC The kringle1 polypeptides (I) are used to screen for agonists and
CC antagonists. Agonists are used to treat subjects in need of enhanced
CC activity or expression of (I). Antagonists are used to treat subjects
CC having need to inhibit the activity or expression of (I). The methods can
CC be used to treat conditions such as cancer, inflammation, autoimmunity,
CC allergy, asthma, rheumatoid arthritis, CNS inflammation, cerebellar
CC degeneration, Alzheimer's disease, Parkinson's disease, multiple
CC sclerosis, amyotrophic lateral sclerosis, head injury damage and other
CC neurological abnormalities, ischemia reperfusion injury, cardiovascular
CC disease, kidney disease, liver disease, ischemic injury, myocardial
CC infarction, hypertension, AIDS, myelodysplastic syndromes
CC and other hematologic abnormalities, aplastic anaemia, male pattern
CC baldness, and bacterial, fungal, protozoan and viral infections. The
CC kringle1 polypeptides may also be used to generate antibodies.
CC Determining the presence or absence of mutations in, and analysing for
CC the presence or absence of expression of, kringle1 polynucleotides can be
CC used to diagnose a disease or susceptibility to a disease related to
CC expression or activity of kringle1 proteins. The polynucleotides may also
CC be used for chromosome identification, and mapping.
XX
XX Sequence 263 AA;
XX
XX Query Match 100.0%; Score 357; DB 20; Length 263;
XX Best Local Similarity 100.0%; Pred. No. 1,9e-34;
XX Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 CFWDNGHL YREDQTS PAPGLRCLNWLDAQSGIAPVSGAGNHSYCRNPDDPRGPMCYV 60
Dd 25 CFWDNGHL YREDQTS PAPGLRCLNWLDAQSGIAPVSGAGNHSYCRNPDDPRGPMCYV 84
XX
XX RESULT 2
XX AAM87769
XX ID AAM87769 standard; Protein; 263 AA.
XX
XX AAM87769;
XX
XX 29-MAR-1999 (first entry)
XX
XX Human tissue plasminogen activator-like protease t-PALP.
XX
XX Tissue plasminogen activator-like protease; t-PALP; human;
XX circulatory system-related disorder; blood clotting; stroke;
XX thrombosis; peripheral arterial occlusion; pulmonary embolism;
XX myocardiothrombosis; diagnosis; therapy.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT Peptide 1..21
XX FT /label= Sig_peptide
XX FT 22..263
XX FT /label= Mat_protein
XX FT 25..84
XX FT /note= "kringle domain"
XX FT Domain 85..263
XX FT /note= "protease domain"
XX FT 22..31
XX FT /note= "epitope-bearing region"
XX FT Peptide 35..44
XX FT /note= "epitope-bearing region"
XX FT Peptide 71..81
XX FT /note= "epitope-bearing region"
XX FT Peptide 91..107

```

```

FT /note= "epitope-bearing region"
FT Peptide 119..128
FT /note= "epitope-bearing region"
FT Peptide 138..147
FT /note= "epitope-bearing region"
FT Peptide 155..167
FT /note= "epitope-bearing region"
FT Peptide 193..203
FT /note= "epitope-bearing region"
FT Peptide 206..215
FT /note= "epitope-bearing region"
FT Peptide 227..237
FT /note= "epitope-bearing region"
FT Peptide 243..252
FT /note= "epitope-bearing region"
XX
XX WO9854199-A1.
XX
XX 03-DEC-1998.
XX
XX 27-MAY-1998; 98WO-US10728.
XX
XX 28-MAY-1997; 97US-0048000.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ebner R, Moore PA, Ruben SM;
XX
XX MPI: 1999-070207/06.
XX
XX N-PSDB; AAV99636.
XX
XX New tissue plasminogen activator-like protease - useful in the
XX diagnosis and treatment of circulatory system-related disorders
XX
XX Claim 1; Page 56-57; 76pp; English.
XX
XX This is the amino acid sequence of tissue plasminogen activator-like
XX protease (t-PALP), a novel member of the serine protease family
XX that shares sequence homology to human tissue plasminogen activator
XX (see AAM87770). The t-PALP sequence was deduced from a cDNA clone
XX (see AAV99636) derived from activated monocytes. The 2.5 kb t-PALP
XX message has also been detected in heart, brain, lung, placenta,
XX liver, skeletal muscle, kidney, pancreas, spleen, thymus, prostate,
XX testis, ovary, small intestine, colon and peripheral blood
XX leukocytes. Isolated nucleic acids encoding amino acids -21 to
XX 242, -20 to 242, 1-242, 4-63 (kringle domain) and 64-242 (protease
XX domain) of t-PALP, or encoding epitope-bearing portions of t-PALP,
XX are also claimed, as are recombinant vectors, host cells, and
XX methods for producing t-PALP polypeptides. t-PALP may be used to
XX detect and treat disorders related to the circulatory system, and
XX to identify agonists and antagonists of t-PALP activity. The
XX homology between t-PALP and tPA indicates that t-PALP may be
XX involved in the regulation of normal and abnormal clotting
XX in e.g. stroke, deep-vein thrombosis, peripheral arterial
XX occlusion, pulmonary embolism and myocardiothrombosis.
XX
XX Sequence 263 AA;
XX
XX Query Match 100.0%; Score 357; DB 20; Length 263;
XX Best Local Similarity 100.0%; Pred. No. 1,9e-34;
XX Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 CFWDNGHL YREDQTS PAPGLRCLNWLDAQSGIAPVSGAGNHSYCRNPDDPRGPMCYV 60
Dd 25 CFWDNGHL YREDQTS PAPGLRCLNWLDAQSGIAPVSGAGNHSYCRNPDDPRGPMCYV 84
XX
XX RESULT 3
XX AAM93748
XX ID AAM93748 standard; Protein; 263 AA.
XX
XX AAM93748;
XX

```

XX	06-NOV-2001	(first entry)	
DE	Human polypeptide, SEQ ID NO: 3727.		
XX			
KW	Human; full length cDNA; cDNA synthesis; oligo-capping.		
XX			
OS	Homo sapiens.		
XX			
FN	EP1130094-A2.		
XX			
PD	05-SEP-2001.		
XX			
PF	07-JUL-2000; 2000EP-0114089.		
XX			
PR	08-JUL-1999; 99JP-0194486.		
PR	11-JAN-2000; 2000JP-0118774.		
XX	02-MAY-2000; 2000JP-0183765.		
PA	(HELI-) HELIX RES INST.		
PI	Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;		
XX	Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;		
XX	WPI: 2001-524255/58.		
DR	N-PSDB; AAK94700.		
PT	830 Primers useful for synthesizing full length cDNA clones and their		
XX	use in genetic manipulation -		
PS	Claim 8; SEQ ID NO 3727; 1380BP + sequence listing; English.		
CC	The invention relates to primers for synthesizing full length cDNA		
CC	clones. 830 cDNA molecules encoding a human protein have been		
CC	isolated and nucleotide sequences of 5' - and 3' -ends of the cDNA		
CC	molecules have been determined. Primers for synthesizing the full length		
CC	cDNA are useful for clarifying the function of the protein encoded by		
CC	the cDNA. The full length clones were obtained by construction of full		
CC	length enriched cDNA libraries that were synthesised by the oligo-capping		
CC	method. The primers enable the production of the full length cDNA easily		
CC	without any special methods. The present sequence is a polypeptide		
CC	encoded by a full length human cDNA of the invention.		
CC	Note: The sequence data for this patent did not form part of the printed		
CC	specification, but was obtained in CD-ROM format directly from EPO.		
XX			
XX	Sequence 263 AA;		
Query Match	100.0%; Score 357; DB 22; Length 263;		
Best Local Similarity	100.0%; Pred. No. 1.9e-34;		
Matches 60; Conservative	100.0%; Mismatches 0; Indels 0; Gaps 0;		
1	CPWDNGHYREDOQTSPAPGLRCLNMLDAQSGLASAPVSGAGNHSYCRNDEDEPRGWCYV 60		
25	CPWDNGHYREDOQTSPAPGLRCLNMLDAQSGLASAPVSGAGNHSYCRNDEDEPRGWCYV 84		
Db			
RESULT 4			
AAE00300			
ID	AAE00300 standard; Protein; 263 AA.		
XX			
AC	AAE00300;		
XX			
DT	13-JUN-2001 (first entry)		
XX			
DE	Human tissue-plasminogen activator-like protease (t-PA/LP).		
XX			
KW	Human; tissue-plasminogen activator-like protease; t-PA/LP;		
KW	therapy; vascular disease; stroke; deep vein thrombosis; keloid; asthma;		
KW	arterial calcification; blood coagulation disorder; cerebroprotective;		
KW	autoimmune system disorder; human immunodeficiency syndrome; cytostatic;		
KW	rheumatoid arthritis; graft-versus-host disease; thyroiditis; candida;		
KW	insulin dependent diabetes; inflammatory eye disease; allergy; antiviral;		
KW	cardiovascular disease; heart disease; arrhythmia; myocardial ischaemia;		
KW	hyperproliferative disorder; hypertrophic scar; neurological disease;		
KW			

KM	Creutzfeldt-Jakob syndrome; neurodegenerative disorder; antiinflammatory
KM	Alzheimer's disease; Parkinson's disease; immunosuppressive; antitungal;
KW	infectious disease; drug screening; gene therapy; neuroprotective;
KW	cancer; ophthalmological; antibacterial; vulnerary.
XX	
OS	Homo sapiens.
XX	
FT	Key
FT	Binding-site
FT	1..165
FT	/note= "Binds to FLAG polypeptide to form
FT	t-PALP-FLAG fusion protein"
FT	Peptide
FT	1..21
FT	/label= Signal_peptide
FT	4..63
FT	/label= Kringle_domain
FT	12..21
FT	/note= "Conserved region"
FT	Protein
FT	22..263
FT	/note= "Human mature tissue-plasminogen activator-like
FT	protease (t-PALP); Binds to FLAG polypeptide to form
FT	t-PALP-FLAG fusion protein"
FT	22..38
FT	/note= "Conserved region"
FT	Region
FT	22..31
FT	/note= "Epitope-bearing portion"
FT	Region
FT	35..44
FT	/note= "Epitope-bearing portion"
FT	Region
FT	39..49
FT	/note= "Conserved region"
FT	50..62
FT	/note= "Conserved region"
FT	63..84
FT	/note= "Conserved region"
FT	64..242
FT	/label= Protease_domain
FT	71..81
FT	/note= "Epitope-bearing portion"
FT	Region
FT	85..97
FT	/note= "Conserved region"
FT	91..107
FT	/note= "Epitope-bearing portion"
FT	Region
FT	100..118
FT	/note= "Conserved region"
FT	119..128
FT	/note= "Epitope-bearing portion"
FT	Region
FT	119..127
FT	/note= "Conserved region"
FT	128..143
FT	/note= "Conserved region"
FT	138..147
FT	/note= "Epitope-bearing portion"
FT	Region
FT	146..163
FT	/note= "Conserved region"
FT	155..167
FT	/note= "Epitope-bearing portion"
FT	Region
FT	164..180
FT	/note= "Conserved region"
FT	186..200
FT	/note= "Conserved region"
FT	193..203
FT	/note= "Epitope-bearing portion"
FT	Region
FT	201..220
FT	/note= "Conserved region"
FT	206..215
FT	/note= "Epitope-bearing portion"
FT	Region
FT	221..236
FT	/note= "Conserved region"
FT	227..237
FT	/note= "Epitope-bearing portion"
FT	Region
FT	237..248
FT	/note= "Conserved region"
FT	243..252
FT	/note= "Epitope-bearing portion"
FT	Region

FT Region 249..263
 FT /note= "Conserved region"
 XX
 XX WO200125252-A1.
 XX
 XX 12-APR-2001.
 XX
 XX 03-OCT-2000; 2000WO-US27239.
 XX
 XX 04-OCT-1999; 99US-0411977.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Moore PA, Ruben SM, Edner R;
 XX WPI; 2001-235402/24.
 XX N-PSDB; AAD03460.
 XX
 XX New (gene encoding and antibody immunospecific for a)
 PT tissue-plasminogen activator-like protease, useful for the diagnosis
 PT and treatment of (cardio)vascular diseases, hyperproliferative
 PT disorders, immune system disorders and cancers -
 XX
 XX Claim 17; Fig 1; 323pp; English.
 XX
 XX The present amino acid sequence is HMS1842 clone human
 CC tissue-plasminogen activator-like protease (t-PA/LP). The t-PA/LP
 CC sequence and their (ant)agonists are useful for the diagnosis and
 CC treatment of vascular diseases e.g. stroke, deep vein thrombosis and
 CC arterial occlusion, blood coagulation disorders, (auto)immune system
 CC disorders e.g. human immunodeficiency syndrome, rheumatoid arthritis,
 CC graft-versus-host disease, thyroiditis, insulin dependent diabetes and
 CC inflammatory eye disease, allergic reactions e.g. asthma, cardiovascular
 CC diseases e.g. heart disease, arrhythmia and myocardial ischaemia,
 CC hyperproliferative disorders, cancers, hypertrophic scars and keloids,
 CC neurological diseases e.g. Creutzfeldt-Jakob syndrome, neurodegenerative
 CC disorders e.g. Alzheimer's disease and Parkinson's disease and infectious
 CC disease e.g. viral, bacterial and fungal infections. The t-PA/LP sequences
 CC are also useful for drug screening. The t-PA/LP nucleotides are useful as
 CC chromosome markers and are involved in gene therapy.
 XX
 XX Sequence 263 AA;
 SQ
 Query Match 100.0%; Score 357; DB 22; Length 263;
 Best Local Similarity 100.0%; Pred. No. 1.9e-34;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CPWDNGHLYREDQTSPPAGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPGRGWCYV 60
 DB 25 CPWDNGHLYREDQTSPPAGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPGRGWCYV 84
 RESULT 5
 AAU86149
 ID AAU86149 standard; Protein; 263 AA.
 XX
 XX AAU86149;
 XX
 XX 15-JUL-2002 (first entry)
 XX
 XX Human PRO264 polypeptide.
 XX
 XX Human; PRO; benign tumour; malignant tumour; lymphoid malignancy;
 KW leukaemia; neuronal disorder; stromal disorder; blastocoeleic disorder;
 KW inflammatory disorder; immune disorder; angiogenic disorder;
 KW cytostatic; neuroprotective.
 XX
 XX Homo sapiens.
 XX
 XX WO200153486-A1.
 XX
 XX 26-JUL-2001.
 XX
 XX

PF 11-FEB-2000; 2000WO-US03565.
 XX
 XX 08-MAR-1999; 99WO-US05028.
 PR 11-MAR-1999; 99US-123972P.
 PR 11-MAY-1999; 99US-133459P.
 PR 02-JUN-1999; 99WO-US12252.
 PR 22-JUN-1999; 99US-140650P.
 PR 22-JUN-1999; 99US-140653P.
 PR 20-JUL-1999; 99US-144758P.
 PR 26-JUL-1999; 99US-146588P.
 PR 28-JUL-1999; 99US-146222P.
 PR 17-AUG-1999; 99US-149395P.
 PR 31-AUG-1999; 99US-151689P.
 PR 01-SEP-1999; 99WO-US20111.
 PR 15-SEP-1999; 99WO-US21099.
 PR 30-NOV-1999; 99WO-US28313.
 PR 01-DEC-1999; 99WO-US28301.
 PR 01-DEC-1999; 99WO-US28634.
 PR 05-JAN-2000; 2000WO-US00219.
 XX
 XX (GENT) GENTECH INC.
 XX
 XX Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AJ, Hillan KJ;
 PI Masters SA, Pan Y, Pitti RM, Roy MA, Smith V, Stone DW;
 PI Watanabe CK, Wood WI;
 XX
 XX WPI; 2002-205567/26.
 XX N-PSDB; ABK40275.
 XX
 XX Thirty five nucleic acids encoding PRO polypeptides, useful for
 PT treating benign or malignant tumours, leukaemias and lymphoid
 PT malignancies, inflammatory, angiogenic and immunologic disorders -
 XX
 XX Claim 61; Fig 44; 302pp; English.
 XX
 XX The present invention relates to the isolation of novel human PRO
 CC polypeptides and the polynucleotide sequences encoding them. The
 CC PRO polypeptides, agonists, antagonists or anti-PRO antibodies are
 CC useful for treating benign or malignant tumours (e.g. renal, kidney,
 CC bladder, breast, etc), leukaemias and lymphoid malignancies, other
 CC disorders such as neuronal, glial, astrocytic, hypothalamic, glandular,
 CC macrophagal, stromal and blastocoeleic disorders, inflammatory, immune
 CC and angiogenic disorders. The polynucleotide sequences are also
 CC useful in gene therapy. AAU86128-AAU86162 represent the human PRO
 CC polypeptides of the invention.
 XX
 XX Sequence 263 AA;
 SQ
 Query Match 100.0%; Score 357; DB 23; Length 263;
 Best Local Similarity 100.0%; Pred. No. 1.9e-34;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CPWDNGHLYREDQTSPPAGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPGRGWCYV 60
 DB 25 CPWDNGHLYREDQTSPPAGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPGRGWCYV 84
 RESULT 6
 AAB43237
 ID AAB43237 standard; Protein; 263 AA.
 XX
 XX AAB43237;
 XX
 XX 08-FEB-2001 (first entry)
 XX
 XX Human ORFX ORF3001 polypeptide sequence SEQ ID NO:6002.
 XX
 XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnerary; antiporiatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopahic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasodilator; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 XX
 XX

KW antihaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX
OS Homo sapiens.
XX
PN MO200058473-A2.
XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000MO-US08621.
XX
PR 31-MAR-1999; 99US-0127607.
XX
PR 02-APR-1999; 99US-0127636.
XX
PR 05-APR-1999; 99US-0127728.
XX
PR 30-MAR-2000; 2000US-0540763.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
PI WPI: 2000-602362/57.
XX
DR N-PSDB; AAC77446.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
PS Claim 11; Page 5181-5182; 5507pp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatocytic; vulnary;
CC antiproliferative; antiparkinsonian; nootropic; neuroprotective;
CC osteoprotective; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antineumatic;
CC antihypertensive; antihypertensive. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ Sequence 263 AA;
Query Match 98.3%; Score 351; DB 21; Length 263;
Best Local Similarity 98.3%; Pred. No. 9.9e-34;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CPWDGHHLYREDQTSFAPGLKCLNWLDAQGLASAPVSGAGNHSYCRNPDEDPGRGWCYV 60
DB 25 CPWDGHHLYREDQTSFAPGLKCLNWLDAQGLASAPVSGAGNHSYCRNPDEDPGRGWCYV 84

RESULT 7
AAV05220
ID AAV05220 standard; Protein: 286 AA.
XX
AC AAV05220;

XX 17-VUN-1999 (first entry)
XX
XX Kringle1 protein sequence.
DE
XX Kringle1; cancer; inflammation; autoimmunity; rheumatoid arthritis; AIDS;
XX CNS inflammation; cerebellar degeneration; Alzheimer's disease; asthma;
XX Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
XX neurological abnormality; ischaemia reperfusion injury; ischaemic injury;
XX cardiovascular disease; kidney disease; liver disease; aplastic anaemia;
XX myocardial infarction; hypertension; allergy; infection;
XX myelodysplastic syndrome; haematologic abnormality; diagnosis; therapy;
XX male pattern baldness.
XX
OS Homo sapiens.
XX
PN WO9911788-A1.
XX
PD 11-MAR-1999.
XX
PF 02-SEP-1998; 98WO-US18270.
XX
PR 01-SEP-1998; 98US-0144869.
XX
PR 02-SEP-1997; 97US-0056032.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Albane EF, Kirkly KK;
XX
PI WPI: 1999-214707/18.
XX
DR N-PSDB; AAX28355.
XX
XX New kringle1 polypeptides and polynucleotides
PT
PT
PS Claim 14; Page 33; 42pp; English.
XX
XX This sequence is a Kringle1 polypeptide of the invention.
CC The kringle1 polypeptides (I) are used to screen for agonists and
CC antagonists. Agonists are used to treat subjects in need of enhanced
CC activity or expression of (I). Antagonists are used to treat subjects
CC having need to inhibit the activity or expression of (I). The methods can
CC be used to treat conditions such as cancer, inflammation, autoimmunity,
CC allergy, asthma, rheumatoid arthritis, CNS inflammation, cerebellar
CC degeneration, Alzheimer's disease, Parkinson's disease, multiple
CC sclerosis, amyotrophic lateral sclerosis, head injury damage and other
CC neurological abnormalities, ischaemia reperfusion injury, cardiovascular
CC disease, kidney disease, liver disease, ischaemic injury, myocardial
CC infarction, hypertension, AIDS, myelodysplastic syndromes
CC and other haematologic abnormalities, aplastic anaemia, male pattern
CC baldness, and bacterial, fungal, protozoan and viral infections. The
CC kringle1 polypeptides may also be used to generate antibodies.
CC Determining the presence or absence of mutations in, and analysing for
CC the presence or absence of expression of, kringle1 polynucleotides can be
CC used to diagnose a disease or susceptibility to a disease related to
CC expression or activity of kringle1 proteins. The polynucleotides may also
CC be used for chromosome identification, and mapping.
XX
SQ Sequence 286 AA;
Query Match 96.1%; Score 343; DB 20; Length 286;
Best Local Similarity 96.7%; Pred. No. 9.7e-33;
Matches 58; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CPWDGHHLYREDQTSFAPGLKCLNWLDAQGLASAPVSGAGNHSYCRNPDEDPGRGWCYV 60
DB 25 CPWDGHHLYREDQTSFAPGLKCLNWLDAQGLASAPVSGAGNHSYCRNPDEDPGRGWCYV 84

RESULT 8
AAV12615
ID AAV12615 standard; Protein: 56 AA.
XX
AC AAV12615;

```

XX 22-JUN-1999 (first entry)
XX
XX Human 5' EST secreted protein SEQ ID NO: 280 from WO 9906553.
DE
XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
XX forensic; gene therapy; chromosome mapping; signal peptide;
XX upstream regulatory sequence; cytokine activity; cell proliferation;
XX differentiation; haematopoiesis regulation; tissue growth regulation;
XX reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
XX thrombolytic; anti-inflammatory; tumour inhibition; antitumour.
XX
XX Homo sapiens.
OS
XX WO9906553-A2.
XX
XX 11-FEB-1999.
XX
XX 31-JUL-1998; 98WO-IB01237.
XX
XX 01-AUG-1997; 97US-0905051.
XX
XX (GEST ) GENSET.
XX
XX Duclert A, Dumas Milne Edwards J, Lacroix B;
XX
XX WPI; 1999-153783/13.
XX
XX N-PSDB; AAX41473.
XX
XX New nucleic acids encoding human secreted proteins - obtained from
XX cDNA libraries derived from umbilical cord, lymph ganglia,
XX lymphocytes and placental tissue
XX
XX Claim 34; Page 376; 411pp; English.
XX
XX AAX41379 to AAX41526 represent 5' expressed sequence tags (ESTs) for
XX human secreted proteins, and encode the proteins given in AAY12521 to
XX AAY12668, respectively. The proteins given represent the signal peptide
XX and an N-terminal fragment of a secreted protein. The nucleic acid
XX sequences can be used for producing secreted human gene products. They
XX can also be used to develop products for diagnosis and therapy. The
XX proteins obtained may have cytokine activity, cell
XX proliferation/differentiation activity, haematopoiesis regulating
XX activity, tissue growth regulating activity, reproductive hormone
XX regulating activity, chemotactic/chemokinetic activity, haemostatic and
XX thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX activity, tumour inhibition activity or other activities. The products
XX can be used in forensic, gene therapy and chromosome mapping procedures.
XX The sequences can also be used for obtaining corresponding promoter
XX sequences. The nucleic acids encoding the signal peptide can be used for
XX directing extracellular secretion of a polypeptide or the insertion of a
XX polypeptide into a membrane, or importing a polypeptide into a cell.
XX
XX Sequence 56 AA;
SQ

```

Query Match 52.1%; Score 186; DB 20; Length 56;
 Best Local Similarity 96.9%; Pred. No. 8.3e-15;
 Matches 31; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFWDNGHLYREDQTSAPAGLRCLNWLDAQSG 32
 Db 25 CFWDNGHLYREDQTSAPAGLRCLNWLDAQNG 56

RESULT 9
 ID AAY12397 standard; Protein; 55 AA.
 AC AAY13397;
 XX
 XX 17-JUN-1999 (first entry)
 DT
 XX Human 5' EST secreted protein SEQ ID NO:428.

```

XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
XX forensic; gene therapy; chromosome mapping; signal peptide;
XX upstream regulatory sequence; cytokine activity; cell proliferation;
XX differentiation; haematopoiesis regulation; tissue growth regulation;
XX reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
XX thrombolytic; anti-inflammatory; tumour inhibition.
XX
XX Homo sapiens.
OS
XX WO9906548-A2.
XX
XX 11-FEB-1999.
XX
XX 31-JUL-1998; 98WO-IB01222.
XX
XX 01-AUG-1997; 97US-0905135.
XX
XX (GEST ) GENSET.
XX
XX Duclert A, Dumas Milne Edwards J, Lacroix B;
XX
XX WPI; 1999-153778/13.
XX
XX N-PSDB; AAX41230.
XX
XX New nucleic acids encoding human secreted proteins - obtained from
XX cDNA libraries prepared from e.g. liver, ovary, brain, prostate,
XX kidney, lung, umbilical cord, placenta and colon tissue
XX
XX Claim 27; Page 744; 824pp; English.
XX
XX AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for
XX human secreted proteins, and encode the proteins given in AAY12261 to
XX AAY12514, respectively. The proteins given represent the signal peptide
XX and an N-terminal fragment of a secreted protein. The nucleic acid
XX sequences can be used for producing secreted human gene products. They
XX can also be used to develop products for diagnosis and therapy. The
XX proteins obtained may have cytokine activity, cell
XX proliferation/differentiation activity, haematopoiesis regulating
XX activity, tissue growth regulating activity, reproductive hormone
XX regulating activity, chemotactic/chemokinetic activity, haemostatic and
XX thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX activity, tumour inhibition activity or other activities. The products
XX can be used in forensic, gene therapy and chromosome mapping procedures.
XX The sequences can also be used for obtaining corresponding promoter
XX sequences. The nucleic acids encoding the signal peptide can be used for
XX directing extracellular secretion of a polypeptide or the insertion of a
XX polypeptide into a membrane, or importing a polypeptide into a cell.
XX
XX Sequence 55 AA;
SQ

```

Query Match 51.8%; Score 185; DB 20; Length 55;
 Best Local Similarity 100.0%; Pred. No. 1.1e-14;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFWDNGHLYREDQTSAPAGLRCLNWLDAQSG 31
 Db 25 CFWDNGHLYREDQTSAPAGLRCLNWLDAQSG 55

RESULT 10
 ID AAW72641 standard; peptide; 39 AA.
 AC AAW72641;
 XX
 XX 05-JAN-1999 (first entry)
 DT
 XX Nervous glia cell growth factor N-terminal peptide #2.
 DE
 XX Nervous glia cell growth factor; human; urine; secretion promoter;
 KM choline acetyltransferase activity enhancer; nervous disease.
 XX


```
OS Homo.sapiens.
XX Key Location/Qualifiers
FH Misc-difference 25 /note= "unspecified"
FT Misc-difference 29 /note= "unspecified"
FT Misc-difference 29 /note= "unspecified"
XX JPI0265498-A.
XX 06-OCT-1998.
XX 24-MAR-1997; 97JP-0090305.
XX 24-MAR-1997; 97JP-0090305.
XX 24-MAR-1997; 97JP-0090305.
XX (NICH-) JAPAN CHEM RES CO LTD.
XX WPI; 1998-589719/50.
XX The present invention describes nervous glia cell growth factor, which
XX is purified from human urine by ultrafiltration, salting-out by ammonium
XX sulphate, gel filtration, ion exchange chromatography and reversed phase
XX chromatography, and has a M.W. of 29 kDa by sodium dodecyl sulphate-
XX polyacrylamide gel electrophoresis. Also described are: (1) a secretion
XX promoter for the nerve growth factor of glia cell consisting of the
XX above growth factor, an enhancer for choline acetyltransferase activity
XX of neuron consisting of the above growth factor; and (2) DNA encoding
XX nervous glia growth factor containing a DNA sequence coding the amino
XX acid sequence shown by the two 39 amino acid sequences as given in Tyr
XX CC AAW72640 and AAW72641, which are identical, except one starts with Tyr
XX CC and the other with Ser (i.e. they are from different DNA transcripts).
XX CC The glia cell growth factor can be prepared in a large amount and the
XX CC factor can be used for the treatment of nervous diseases.
XX Sequence 39 AA;
XX
XX Query Match 51.3%; Score 183; DB 19; Length 39;
XX Best Local Similarity 94.3%; Pred. No. 1.3e-14;
XX Matches 33; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 FWDNGHLYREDQTSPPAGLRCLNWLDAQSGLASAP 36
DB 5 FWDNGHLYREDQTSPPAGLRCLNWLDAQSGLASAP 39
RESULT 11
AAW72640
ID AAW72640 standard; peptide; 39 AA.
XX
XX AAW72640;
XX
XX 05-JAN-1999 (first entry)
XX
XX Nervous glia cell growth factor N-terminal peptide #1.
XX
XX Nervous glia cell growth factor; human; urine; secretion promoter;
XX choline acetyltransferase activity enhancer; nervous disease.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FH Misc-difference 25 /note= "unspecified"
XX FT Misc-difference 29 /note= "unspecified"
XX FT Misc-difference 29 /note= "unspecified"
XX JPI0265498-A.
```

```
XX
XX 06-OCT-1998.
XX 24-MAR-1997; 97JP-0090305.
XX 24-MAR-1997; 97JP-0090305.
XX 24-MAR-1997; 97JP-0090305.
XX (NICH-) JAPAN CHEM RES CO LTD.
XX WPI; 1998-589719/50.
XX Nervous glia cell growth factor derived from human urine - used for
XX treatment of nervous diseases
XX
XX Claim 2; Fig 6; 14pp; Japanese.
XX
XX The present invention describes nervous glia cell growth factor, which
XX is purified from human urine by ultrafiltration, salting-out by ammonium
XX sulphate, gel filtration, ion exchange chromatography and reversed phase
XX chromatography, and has a M.W. of 29 kDa by sodium dodecyl sulphate-
XX polyacrylamide gel electrophoresis. Also described are: (1) a secretion
XX promoter for the nerve growth factor of glia cell consisting of the
XX above growth factor, an enhancer for choline acetyltransferase activity
XX of neuron consisting of the above growth factor; and (2) DNA encoding
XX nervous glia growth factor containing a DNA sequence coding the amino
XX acid sequence shown by the two 39 amino acid sequences as given in Tyr
XX CC AAW72640 and AAW72641, which are identical, except one starts with Tyr
XX CC and the other with Ser (i.e. they are from different DNA transcripts).
XX CC The glia cell growth factor can be prepared in a large amount and the
XX CC factor can be used for the treatment of nervous diseases.
XX Sequence 39 AA;
XX
XX Query Match 51.3%; Score 183; DB 19; Length 39;
XX Best Local Similarity 94.3%; Pred. No. 1.3e-14;
XX Matches 33; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 FWDNGHLYREDQTSPPAGLRCLNWLDAQSGLASAP 36
DB 5 FWDNGHLYREDQTSPPAGLRCLNWLDAQSGLASAP 39
RESULT 12
AAW54154
ID AAW54154 standard; protein; 527 AA.
XX
XX AAW54154;
XX
XX 20-JUL-1998 (first entry)
XX
XX t-PA mutant (N142S).
XX
XX Amino acid substitution; t-PA; vascular disorder; prevention;
XX fibrin deposition; adhesion formation.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX FH Domain 1..44 /note= "Finger domain"
XX FT Domain 45..91 /note= "Growth factor domain"
XX FT Domain 92..173 /note= "Kring1-1 domain"
XX FT Domain 180..261 /note= "Kring1-2 domain"
XX FT Domain 264..527 /note= "Serine protease domain"
XX misc_difference 142 /note= "N changed from wt to S in mutant"
XX
XX US5736135-A.
```

```
PD 07-APR-1998.
XX
XX 13-FEB-1995; 95US-0389615.
XX
PR 11-JUL-1991; 91US-0728456.
PR 26-JAN-1993; 93US-0008940.
PR 01-APR-1994; 94US-0221660.
PR 13-FEB-1995; 95US-0389615.
XX
XX (GETH ) GENENTECH INC.
XX
XX Goeddel DV, Leung DMH, Rice GC;
XX
XX WPI; 1998-239153/21.
XX
XX Mutant tissue plasminogen activator proteins - useful for treating
XX PT vascular disorders, preventing tissue adhesion(s), etc.
XX
XX
XX PS Claim 6; Page -; 24pp; English.
XX
XX CC Mutant tissue plasminogen activator proteins (AAW54147-W54158). are
XX CC created by single or multiple amino acid substitutions. Compositions
XX CC containing the t-PA variant are used for treating vascular disorders, for
XX CC preventing fibrin deposition or for preventing adhesion formation or
XX CC reformation. Note: This sequence is not given in the specification but
XX CC was created from the wild type by the indexer.
XX
XX SQ Sequence 527 AA;

Query Match 41.7%; Score 149; DB 19; Length 527;
Best Local Similarity 44.9%; Pred. No. 2,4e-09;
Matches 31; Conservative 4; Mismatches 22; Indels 12; Gaps 3;

QY 1 CFWDNGHLVREQTSPPAGLRCLNWLDAQSGLASAPV-----GAGNHSYCRNPDE 51
DB 92 CYEDQGISIRGTWSTAESGAECTNW--NSSALAKPKYSGRRPDARLGLGNHSYCRNPDR 149
QY 52 DPGGPMCYV 60
DB 150 DSK-PWCYV 157

RESULT 13
AAW54157
ID AAW54157 standard; protein; 527 AA.
XX
XX AAW54157;
XX
XX 20-JUL-1998 (first entry)
XX
XX t-PA mutant (N142S).
XX
XX DE Amino acid substitution; t-PA; vascular disorder; prevention;
XX KW fibrin deposition; adhesion formation.
XX
XX OS Synthetic.
XX
XX Key Location/Qualifiers
XX FH 1..44
XX FT /note= "Finger domain"
XX FT 45..91
XX FT /note= "Growth factor domain"
XX FT 92..173
XX FT /note= "Kring1e-1 domain"
XX FT 180..261
XX FT /note= "Kring1e-2 domain"
XX FT 264..527
XX FT /note= "Serine protease domain"
XX FT misc_difference 142
XX FT /note= "N changed from wt to S in mutant"
XX
XX PN US5736135_A.
```

```
PD 07-APR-1998.
XX
XX 13-FEB-1995; 95US-0389615.
XX
XX 11-JUL-1991; 91US-0728456.
XX 26-JAN-1993; 93US-0008940.
XX 01-APR-1994; 94US-0221660.
XX 13-FEB-1995; 95US-0389615.
XX
XX (GETH ) GENENTECH INC.
XX
XX Goeddel DV, Leung DMH, Rice GC;
XX
XX WPI; 1998-239153/21.
XX
XX Mutant tissue plasminogen activator proteins - useful for treating
XX PT vascular disorders, preventing tissue adhesion(s), etc.
XX
XX
XX PS Claim 2; Page -; 24pp; English.
XX
XX CC Mutant tissue plasminogen activator proteins (AAW54147-W54158) are
XX CC created by single or multiple amino acid substitutions. Compositions
XX CC containing the t-PA variant are used for treating vascular disorders, for
XX CC preventing fibrin deposition or for preventing adhesion formation or
XX CC reformation. Note: This sequence is not given in the specification but
XX CC was created from the wild type by the indexer.
XX
XX SQ Sequence 527 AA;

Query Match 41.7%; Score 149; DB 19; Length 527;
Best Local Similarity 44.9%; Pred. No. 2,4e-09;
Matches 31; Conservative 4; Mismatches 22; Indels 12; Gaps 3;

QY 1 CFWDNGHLVREQTSPPAGLRCLNWLDAQSGLASAPV-----GAGNHSYCRNPDE 51
DB 92 CYEDQGISIRGTWSTAESGAECTNW--NSSALAKPKYSGRRPDARLGLGNHSYCRNPDR 149
QY 52 DPGGPMCYV 60
DB 150 DSK-PWCYV 157

RESULT 14
AAP82587
ID AAP82587 standard; protein; 390 AA.
XX
XX AAP82587;
XX
XX 03-NOV-1990 (first entry)
XX
XX DE Modified tissue plasminogen activator lacking F and G region and
XX DE kring1e region 2 and Q-96, I-98 and S-119 subseq for N, T and W resp.
XX
XX KW Modified tissue plasminogen activator; tPA; thrombosis;
XX KW N-glycosylation.
XX
XX PN JP63230084-A.
XX
XX 26-SEP-1988.
XX
XX 20-MAR-1987; 87JP-0064340.
XX
XX 20-MAR-1987; 87JP-0064339.
XX
XX (EISA ) EISA KK.
XX
XX WPI; 1988-311962/44.
XX
XX N-PSDB; AAP82107.
XX
XX Multiple modified tissue plasminogen activator - which lacks F and G
XX PT regions, is useful in treatment of thrombosis and has improved half life.
XX
XX PS Disclosure; ); 29pp; Japanese.
```

CX		* Plasmaid encoding the modified tPA is 99-2820 and its transformant is E.coli RRI-Zem 99-2820 (PERM P-9125).
CC		This modified tPA, used to treat thromboisis, is of high quality and has a longer half life period in blood.
CC		See also AANB1492, AANB2183-NB2188 and J6323J0083.
SQ	Sequence	390 AA;
OY	Query Match	41.5%; Score 148; DB 9; Length 390; Best Local Similarity 43.5%; Pred. No. 2.3e-09;
Dd	Matches	30; Conservative 5; Mismatches 22; Indels 12; Gaps 3.
Oy	1 CPMDNGHLYREDQTSPPAPGLRCINLWIDAOGLASHPVS-----GAGNHSYCRNPDE 51 Db 43 CYPDNDSTSRKGTGSTAESSGAECTNN--NSMALAQFYSGRRPDAIRLGAGNHNYCRNPDR 100	
Oy	52 DPRGPWCYV 60 Dd : 101 DSK-PWCYV 108	
RESULT 15	AAR70889 standard; Protein; 483 AA. AAR70889	
XX	AAR70889;	
AC	05-OCT-1995 (first entry)	
XX		
XX	Human tissue PA variant (deltaI-44,N103,R252,E275,Q277) .	
XX		
KM	Human wild type tissue plasminogen activator; fibrinolytic; variant; triptetide; glycosylation site; half-life; clearance rate; blood; clot-dissolving agent; vascular disease; thrombosis; artery; vein.	
KW		
OS	Synthetic.	
XX		
FH	Key	Location/Qualifiers
FT	Domain	1..47
FT	Domain	/label= growth factor domain
FT	Domain	48..129
FT	Domain	/label= kringle 1 domain
FT	Domain	136..217
FT	Domain	/label= kringle 2 domain
FT	Domain	220..483
FT	Disulfide-bond	7..18
FT	Disulfide-bond	12..29
FT	Disulfide-bond	31..40
FT	Disulfide-bond	48..129
FT	Disulfide-bond	169..111
FT	Disulfide-bond	100..124
FT	Disulfide-bond	136..217
FT	Disulfide-bond	157..199
FT	Disulfide-bond	288..212
FT	Disulfide-bond	220..351
FT	Disulfide-bond	263..379
FT	Disulfide-bond	271..340
FT	Disulfide-bond	365..440
FT	Disulfide-bond	397..413
FT	Disulfide-bond	430..448
FT	Modified-site	73
FT	Modified-site	/label= N-linked glycosylation site
FT	Modified-site	140
FT	Modified-site	/label= N-linked glycosylation site
FT	Modified-site	404
FT	Misc-difference	/label= N-linked glycosylation site
FT	233	
FT	/label= OTHER	
FT	/note= "represented in specification as o"	
PV	US5385732-A.	

```

XX 31-JAN-1995.
XX PD
XX 20-MAY-1988; 88US-0196909.
XX PF
XX 20-MAY-1988; 88US-0196909.
XX PR
XX 15-FEB-1990; 90US-0480691.
XX PR
XX 21-JUN-1992; 92US-0824740.
XX PR
XX 22-MAR-1993; 93US-0035427.
XX PR
XX (GENT) GENENTECH INC.
XX PA
XX Anderson S, Brady KM, Keyt BA, Presta LG;
XX WPI; 1995-081536/11.
XX DR
XX
XX New tissue plasminogen activator variants - having an N-linked
XX PT tripeptidyl glycosylation sequence inserted to increase plasma
XX PT half-life
XX PS
XX Claim 12; ; 34pp; English.
XX
XX The amino acid sequence of the human tissue plasminogen activator
XX CC (t-PA) variant delc1-44,M103,R252,E275,Q277 (sic)
XX CC This sequence varies from the wild type sequence (AAR70842) by:
XX CC (a) a deletion of amino acids (AA) 1-44,
XX CC (b) substitution of AA at pos.: 103 - G to N, 252 - T to R, 275 - R to
XX CC E; 277 - K to O.
XX CC The numbers correspond to the residue positions in the wild type t-PA.
XX CC This sequence is one of a series of fibrinolytically active variant
XX CC t-PAs (see AAR70843-70908 + AAR79144) - the variants are modified to
XX CC contain one or more amino acid substitutions, which provide an
XX CC bonyl-Ser/Thr-tripeptidyl sequence starting at the positions 57-61,
XX CC 63-66, 89, 101, 103-105, 106, 107, 109,112, or 250 of the wild type t-PA
XX CC amino acid sequence. The Aaa of the tripeptidyl sequence contains an
XX CC N-linked glycosylation site. The variant t-PAs exhibit fibrinolytic
XX CC activity and have longer half-lives and slower clearance rates from the
XX CC blood as compared to native t-PA. The variants can be used as
XX CC clot-dissolving agents in the treatment of vascular diseases or
XX CC conditions such as deep vein thrombosis or peripheral arterial
XX CC thrombosis.
XX CC
XX
XX Sequence 483 AA;
XX SQ
XX
XX Query Match 41.5%; Score 148; DB 16; Length 483;
XX Best Local Similarity 43.5%; Pred. No. 2.9e-09;
XX Matches 30; Conservative 6; Mismatches 21; Indels 12; Gaps 3;
XX
XX 1 CPMNDNHLYREDOQSPAPGLRLNLTNDQSGIAAPVS-----GAGNHSYCRNPDE 51
XX Db 48 CYPDQGISIRGWSMTASGAEFCNM--NSSALAKPYSGRPAIRLIGAHNNYCRNPDR 105
XX
XX 52 DRRGPGWCV 60
XX Db 106 DSK-PWCYV 113

```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:19:28 ; Search time 4.24116 Seconds
(without alignments)
864.895 Million cell updates/sec

Title: US-10-057-951-2_COPY_25_84

Perfect score: 1 CFWDNGHLYREDQTSAPAGL.....GNHSYCRNPDEDRGWCYV 60

Sequence: BLOSUM62

Scoring table: Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications, AA:*

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PC7_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PC7US_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	357	100.0	263	9 US-10-102-704-2	Sequence 2, Appl1
2	357	100.0	263	9 US-10-057-951-2	Sequence 2, Appl1
3	357	100.0	263	10 US-09-084-491A-2	Sequence 2, Appl1
4	146	40.9	527	9 US-09-987-457-18	Sequence 18, Appl1
5	146	40.9	527	9 US-09-987-455-19	Sequence 19, Appl1
6	146	40.9	562	9 US-09-974-298-145	Sequence 145, Appl1
7	146	40.9	562	10 US-09-969-271-7	Sequence 7, Appl1
8	131	36.7	88	10 US-09-880-503-1	Sequence 9, Appl1
9	131	36.7	96	10 US-09-880-503-9	Sequence 9, Appl1
10	131	36.7	135	10 US-09-880-503-4	Sequence 4, Appl1
11	131	36.7	138	9 US-10-237-667-12	Sequence 12, Appl1
12	131	36.7	138	9 US-10-237-708-12	Sequence 12, Appl1
13	131	36.7	138	9 US-10-237-866-12	Sequence 12, Appl1
14	131	36.7	138	9 US-10-237-871-12	Sequence 12, Appl1
15	131	36.7	138	10 US-09-984-186-12	Sequence 12, Appl1
16	131	36.7	143	10 US-09-880-503-8	Sequence 8, Appl1
17	131	36.7	403	10 US-09-880-503-6	Sequence 6, Appl1
18	131	36.7	411	10 US-09-880-503-3	Sequence 3, Appl1
19	131	36.7	431	9 US-10-076-421-2	Sequence 2, Appl1

20	131	36.7	431	10 US-09-264-468B-1	Sequence 1, Appl1
21	127	35.6	326	9 US-10-057-951-3	Sequence 3, Appl1
22	127	35.6	354	9 US-09-987-457-10	Sequence 10, Appl1
23	127	35.6	354	9 US-09-987-455-11	Sequence 11, Appl1
24	127	35.6	372	9 US-10-102-704-3	Sequence 3, Appl1
25	127	35.6	372	10 US-09-084-491A-3	Sequence 3, Appl1
26	127	35.6	377	9 US-09-987-455-8	Sequence 8, Appl1
27	121	33.9	79	9 US-09-335-325-8	Sequence 8, Appl1
28	121	33.9	79	10 US-10-131-241-8	Sequence 2, Appl1
29	121	33.9	79	10 US-09-753-064-2	Sequence 8, Appl1
30	121	33.9	79	10 US-09-761-120-8	Sequence 8, Appl1
31	121	33.9	160	9 US-09-335-325-35	Sequence 35, Appl1
32	121	33.9	160	9 US-10-131-241-35	Sequence 35, Appl1
33	121	33.9	160	10 US-09-761-120-35	Sequence 35, Appl1
34	121	33.9	250	9 US-09-335-325-30	Sequence 30, Appl1
35	121	33.9	250	9 US-10-131-241-30	Sequence 30, Appl1
36	121	33.9	250	10 US-09-761-120-30	Sequence 30, Appl1
37	121	33.9	260	9 US-10-131-241-61	Sequence 61, Appl1
38	121	33.9	339	9 US-09-335-325-3	Sequence 3, Appl1
39	121	33.9	339	9 US-10-131-241-3	Sequence 3, Appl1
40	121	33.9	339	10 US-09-788-142-3	Sequence 3, Appl1
41	121	33.9	339	10 US-09-761-120-3	Sequence 3, Appl1
42	121	33.9	352	9 US-09-335-325-40	Sequence 40, Appl1
43	121	33.9	352	9 US-10-131-241-40	Sequence 40, Appl1
44	121	33.9	352	10 US-09-761-120-40	Sequence 40, Appl1
45	121	33.9	368	10 US-09-761-120-42	Sequence 42, Appl1

ALIGNMENTS

RESULT 1
US-10-102-704-2
; Sequence 2, Application US/10102704
; Patent No. US20020164768A1
; GENERAL INFORMATION:
; APPLICANT: Moore et al.
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protein
; FILE REFERENCE: PF378C1
; CURRENT APPLICATION NUMBER: US/10/102,704
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/084,491
; PRIOR FILING DATE: 1998-05-27
; PRIOR APPLICATION NUMBER: 60/048,000
; PRIOR FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-102-704-2

Query Match 100.0%; Score 357; DB 9; Length 263;
Best Local Similarity 100.0%; Pred. No. 4.9e-32;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFWDNGHLYREDQTSAPAGLRCINMLDAOSGLASAPVGAGNHSYCRNPDEDRGWCYV 60
DB 25 CFWDNGHLYREDQTSAPAGLRCINMLDAOSGLASAPVGAGNHSYCRNPDEDRGWCYV 84

RESULT 2
US-10-057-951-2
; Sequence 2, Application US/10057951
; Patent No. US20020177213A1
; GENERAL INFORMATION:
; APPLICANT: Moore et al.
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
; FILE REFERENCE: PF378P1
; CURRENT APPLICATION NUMBER: US/10/057,951
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US 09/411,977

;; PRIOR FILING DATE: 1999-10-04
;; PRIOR APPLICATION NUMBER: US 09/084,491
;; PRIOR FILING DATE: 1998-05-27
;; PRIOR APPLICATION NUMBER: US 60/048,000
;; PRIOR FILING DATE: 1997-05-28
;; NUMBER OF SEQ ID NOS: 31
;; SOFTWARE: PatentIn Ver. 3.1
;; SEQ ID NO 2
;; LENGTH: 263
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-057-951-2

Query Match 100.0%; Score 357; DB 9; Length 263;
Best Local Similarity 100.0%; Pred. No. 4.9e-32;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFWDNGLYREDQTSAPGLRCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGRCVCY 60
Db 25 CFWDNGLYREDQTSAPGLRCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGRCVCY 84

RESULT 3
US-09-084-491A-2

;; Sequence 2, Application US/09084491A
;; Patent No. US2002061576A1
;; GENERAL INFORMATION:
;; APPLICANT: MOORE, PAUL A.
;; APPLICANT: RUBEN, STEVEN M.
;; APPLICANT: BENER, REINHARD
;; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
;; NUMBER OF SEQUENCES: 16
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
;; STREET: 9410 KEY WEST AVENUE
;; CITY: ROCKVILLE
;; STATE: MD
;; COUNTRY: USA
;; ZIP: 20850
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/084,491A
;; FILING DATE: 27-MAY-1998
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BROOKES, ANDERS A.
;; REGISTRATION NUMBER: 36,373
;; REFERENCE/DOCKET NUMBER: PF378
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (301) 309-8504
;; TELEFAX: (301) 309-8439
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 263 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-09-084-491A-2

Query Match 100.0%; Score 357; DB 10; Length 263;
Best Local Similarity 100.0%; Pred. No. 4.9e-32;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFWDNGLYREDQTSAPGLRCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGRCVCY 60
Db 25 CFWDNGLYREDQTSAPGLRCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGRCVCY 84

RESULT 4

US-09-987-457-18
;; Sequence 18, Application US/09987457
;; Publication No. US20030013150A1
;; GENERAL INFORMATION:
;; APPLICANT: Manosroi, Aranya
;; APPLICANT: Manosroi, Jiradej
;; APPLICANT: Tayapiwatana, Chatchai
;; APPLICANT: Goetz, Friedrich
;; APPLICANT: Werner, Rolf-Guenther
;; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
;; FILE REFERENCE: 0652.2180001
;; CURRENT APPLICATION NUMBER: US/09/987,457
;; PRIOR FILING DATE: 2001-11-14
;; PRIOR APPLICATION NUMBER: 60/268,573
;; PRIOR FILING DATE: 2001-02-15
;; PRIOR APPLICATION NUMBER: GB 00 27 782.2
;; PRIOR FILING DATE: 2000-11-14
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 18
;; LENGTH: 527
;; TYPE: PRT
;; ORGANISM: Homo sapiens (CPA)
US-09-987-457-18

Query Match 40.9%; Score 146; DB 9; Length 527;
Best Local Similarity 43.5%; Pred. No. 1.9e-08;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

Qy 1 CFWDNGLYREDQTSAPGLRCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGRCVCY 51
Db 92 CYEDGISTYRGTWSTASGAECTNM--NSSALAKKYSRRRDALRLGNNHYCRNPDR 149

Qy 52 DPGRCVCY 60
Db 150 DSK-PWCYV 157

RESULT 5
US-09-987-455-19
;; Sequence 19, Application US/09987455
;; Publication No. US20030049729A1
;; GENERAL INFORMATION:
;; APPLICANT: Aranya Manosroi
;; APPLICANT: Jiradej Manosroi
;; APPLICANT: Chatchai Tayapiwatana
;; APPLICANT: Friedrich Goetz
;; APPLICANT: Rolf-Guenther Werner
;; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
;; TITLE OF INVENTION: DNA-Derived tPA or K25 Molecules
;; FILE REFERENCE: 0652.2190001
;; CURRENT APPLICATION NUMBER: US/09/987,455
;; CURRENT FILING DATE: 2001-11-14
;; PRIOR APPLICATION NUMBER: 60/268,574
;; PRIOR FILING DATE: 2001-02-15
;; PRIOR APPLICATION NUMBER: GB 0027779.8
;; PRIOR FILING DATE: 2000-11-14
;; NUMBER OF SEQ ID NOS: 25
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 19
;; LENGTH: 527
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-987-455-19

Query Match 40.9%; Score 146; DB 9; Length 527;
Best Local Similarity 43.5%; Pred. No. 1.9e-08;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

Qy 1 CFWDNGLYREDQTSAPGLRCLNWLDAOSGLASAPVSGAGNHSYCRNPDEDPGRCVCY 51
Db 92 CYEDGISTYRGTWSTASGAECTNM--NSSALAKKYSRRPDALRLGNNHYCRNPDR 149

```
QY      52 DFRGPMCYV 60
          | : |||||
Db      150 DSK-PMCYV 157
```

RESULT 6

```

US-09-974-298-145
/ Sequence 145, Application US/09974298
/ Patent No. US20020156263A1
/ GENERAL INFORMATION:
/ APPLICANT: Chen, Hwei-Mei
/ TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
/ FILE REFERENCE: PA-0037 P
/ CURRENT APPLICATION NUMBER: US/09/974,298
/ CURRENT FILING DATE: 2001-10-04
/ PRIOR APPLICATION NUMBER: 60/238,331
/ PRIOR FILING DATE: 2000-05-10
/ NUMBER OF SEQ. ID NOS: 194
/ SOFTWARE: PERL Program
/ SEQ. ID NO 145
/ LENGTH: 562
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ OTHER INFORMATION: Incyte ID No. US20020156263A1 1001470CD1
US-09-974-298-145

```

Query Match	40.98;	Score 146;	DB 9;	Length 562;
-------------	--------	------------	-------	-------------

DQ 1 CFMNGHLTKREDQTSAPGLRCINLMDAQSGLASAPV-----GAGNHSYCRNPDE 51
| : | | | : | | | | | : | | | |
Db 127 CYEDQGISYRGTWSTAESGAECTNM--NSSALMQKPYSGRRPDATRLGLGNHNYCRNPDR 184

QY	52	DPRGWCYV	60
		{ :	
Db	185	DSK-PWCYV	192

RESULT 7

```

US-09-969-271-7
; Sequence 7, Application US/09969271
; Patent No. US20020098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB))
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 562
; TYPE: PR1
; ORGANISM: Homo sapiens
; US-09-969-271-7

```

Query Match 40.9%; Score 146; DB 10; Length 562;

OY 1 CFFDNGHLRYEDQTPAPGLRCLNWLDAQSGLASAPVS-----GAGHSYCNPDE 51
 | : | | | | | | | | | | | | | | | |
Dd 127 CYEDQGISYRGTWSTAESGAECTNW--NSSALAQKPYSGRRPDALRLGLGNHNYCRNPDR 184
 | : | | | | | | | | | | | | | | | |

Qy 52 DPRGPWCYV 60
| : |||||
Db 185 DSK-PWCYV 192

RESULT 8
IIS-09-88

Sequence 1, Application US/09880503
Patent No. US20020131964A1
GENERAL INFORMATION:
APPLICANT: CINES, Douglas B
APPLICANT: HIGAZI, Abd Al-Roof
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
FILE REFERENCE: 9596-331
CURRENT APPLICATION NUMBER: US/09/880,503
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/212,847
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 88
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-503-1

Query Match	36.7%;	Score 131;	DB 10;	Length 88;
Best Local Similarity	40.3%;	Pred. No. 1.3e-07;		
Matches 27; Conservative	7;	Mismatches 25;	Indels 8;	Gaps 2

QY 1 CFWDNGHLREDDQTSAPRGRLCNWLDA-----QSGLASAPVSGAGNHSYCRANDEDP 533

Db 3 CYEGNGHFFYRGKASTDMGRPCLEPWN SATVLQQTVAHARSDALQGLGKANYCRNPD-NR 61

QY	54	RGPWCYV	60
Db	62	RRPWCYV	68

RESULT 9

```

US-09-880-503-9
; Sequence 9, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGARTI, Abd Al-Rooof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880.503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 9
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-880-503-9

```

Query Match	36.7%;	Score 131;	DB 10;	Length 96;
Best Local Similarity	40.3%;	Pred. No. 1.5e-07;		
Matches 27; Conservative	7;	Mismatches 25;	Indels 8;	Gaps 2;

QY 1 CFMDNGHLTYREDQTSAPGLRCLNWLDA-----QSGLASAPVSGAGNHSYCRNPDEDF 53

Db 3 CYEGNGHFFRGKASTDTMGRPCLPWNSATVLQQTTHAHRSDALQLGGLKENVCRNPD-NR 61

QY	54	RGPWCYV	60
Db	62	RRPWCYV	68

RESULT 10

TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-708-12

Query Match 36.7%; Score 131; DB 9; Length 138;
Best Local Similarity 40.3%; Pred. No. 2.2e-07;
Matches 27; Conservative 7; Mismatches 25; Indels 8; Gaps 2;

Qy 1 CFWNGHLYREDQTSAPARCLNWLDA-----OSGLASAPVSGAGNHSYCRNPDEDP 53
Db 53 CYEGNGHYRGKASTTWGRPCLPWNSATVLOQTYHAHRSDLQLGLGKHNYCRNPD-NR 111
Qy 54 RGPWCYV 60
Db 112 RRPWCYV 118

RESULT 13
US-10-237-866-12
Sequence 12, Application US/10237866
Publication No. US20030036171A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcoia Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,866
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-866-12

Query Match 36.7%; Score 131; DB 9; Length 138;
Best Local Similarity 40.3%; Pred. No. 2.2e-07;
Matches 27; Conservative 7; Mismatches 25; Indels 8; Gaps 2;

Qy 1 CFWNGHLYREDQTSAPARCLNWLDA-----OSGLASAPVSGAGNHSYCRNPDEDP 53
Db 53 CYEGNGHYRGKASTTWGRPCLPWNSATVLOQTYHAHRSDDLGLGKHNYCRNPD-NR 111
Qy 54 RGPWCYV 60
Db 112 RRPWCYV 118

RESULT 14
US-10-237-871-12
Sequence 12, Application US/10237871
Publication No. US20030036172A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcoia Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,871
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-871-12

```
Search completed: April 7, 2003, 09:31:28
Job time : 4.2416 secs
```

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:17:08 ; Search time 51.2682 Seconds
(without alignments)
754.542 Million cell updates/sec

Title: US-10-057-951-2_COPY_25_84
Perfect score: 357
Sequence: 1 CFMNGHLYREDQTSAPGL.....GNHSYCRNPDEDPGWCYV 60

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 64473110*residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Pending Patents_AA Main:*

- 1: /cgn2_6/ptodata/1/paa/PCRTUS_COMB.pep.*
- 2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
- 3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
- 4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
- 5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
- 6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
- 7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
- 8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
- 9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
- 10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
- 11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
- 12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
- 13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
- 14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
- 15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
- 16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
- 17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
- 18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
- 19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
- 20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
- 21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
- 22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
- 23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
- 24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
- 25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
- 26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
- 27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	357	100.0	109	22	US-09-834-366-13839
2	357	100.0	109	27	US-60-197-873-13839
3	357	100.0	117	1	PCT-US01-14827-8897
4	357	100.0	126	17	US-09-307-140-979
5	357	100.0	126	22	US-09-817-076-979
6	357	100.0	146	1	PCT-US02-05301-237

7	357	100.0	146	1	PCT-US02-05301-237	Sequence 311, App
8	357	100.0	178	27	US-60-213-800-237	Sequence 237, App
9	357	100.0	250	14	US-09-001-403-64	Sequence 64, App
10	357	100.0	263	1	PCT-US00-27239-2	Sequence 2, App
11	357	100.0	263	1	PCT-US02-05301-164	Sequence 164, App
12	357	100.0	263	1	PCT-US98-10728-2	Sequence 2, App
13	357	100.0	263	14	US-09-084-491A-2	Sequence 2, App
14	357	100.0	263	14	US-09-144-889-2	Sequence 2, App
15	357	100.0	263	20	US-09-611-526-3727	Sequence 3727, App
16	357	100.0	263	21	US-09-791-537-39562	Sequence 39562, App
17	357	100.0	263	23	US-09-927-786-44	Sequence 44, App
18	357	100.0	263	24	US-10-057-951-2	Sequence 2, App
19	357	100.0	263	25	US-10-102-704-2	Sequence 2, App
20	357	100.0	263	25	US-10-210-951-44	Sequence 44, App
21	357	100.0	263	26	US-10-211-884-44	Sequence 44, App
22	357	100.0	263	26	US-10-211-884-44	Sequence 44, App
23	357	100.0	263	27	US-60-048-000-2	Sequence 2, App
24	357	100.0	263	27	US-60-230-435-1580	Sequence 1580, App
25	357	100.0	263	27	US-60-389-987-244	Sequence 244, App
26	357	100.0	263	27	US-60-412-418-244	Sequence 244, App
27	357	100.0	308	27	US-60-207-315-471	Sequence 471, App
28	354	99.2	295	1	PCT-US01-14827-8900	Sequence 8900, App
29	343	96.1	286	15	US-09-144-889-4	Sequence 4, App
30	306	85.7	257	1	PCT-US01-14827-8898	Sequence 8898, App
31	186	52.1	56	13	US-08-905-051-280	Sequence 280, App
32	186	52.1	56	19	US-08-905-135-428	Sequence 2338, App
33	185	51.8	55	13	US-08-905-135-428	Sequence 428, App
34	185	51.8	55	19	US-09-547-598C-2126	Sequence 2126, App
35	154	43.1	566	21	US-09-791-537-140530	Sequence 140530, App
36	149	41.7	527	18	US-09-422-601-1	Sequence 1, App
37	146	40.9	82	3	US-07-929-181B-44	Sequence 44, App
38	146	40.9	399	27	US-60-208-020-156	Sequence 156, App
39	146	40.9	399	27	US-60-209-043-206	Sequence 206, App
40	146	40.9	472	13	US-08-984-462-2	Sequence 2, App
41	146	40.9	472	13	US-08-984-462-2	Sequence 2, App
42	146	40.9	472	14	US-09-032-861-12	Sequence 12, App
43	146	40.9	472	14	US-09-038-002-2	Sequence 2, App
44	146	40.9	472	14	US-09-067-929-12	Sequence 12, App
45	146	40.9	472	14	US-09-079-343-2	Sequence 2, App

ALIGNMENTS

US-09-834-366-13839
RESULT 1
Sequence 13839, Application US/09834366
GENERAL INFORMATION:
APPLICANT: Bejani, Stephane
APPLICANT: Tanaka, Hiroaki
APPLICANT: Dumas Milne Edwards, Jean Baptiste
APPLICANT: Jobert, Severin
APPLICANT: Giordano, Jean-Yves
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: 81.US2.REG
CURRENT APPLICATION NUMBER: US/09/834,366
CURRENT FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: US 60/197,873
PRIOR FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 52153
SOFTWARE: Patent.pm
SEQ ID NO 13839
LENGTH: 109
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -21...-1
US-09-834-366-13839
Query Match 100.0%; Score 357; DB 22;
Best Local Similarity 100.0%; Pred. No. 8, 1e-34;
Matches 60; Conservative 0; Mismatches 0; Indels 0;

QY 1 CFWDNGHLTYREDQTSAPAGRLCLNWLDAQSLASAPVSGAGNHSYCRNPDEDPRGWCYV 600

Db 25 CFWDNGHLTYREDQTSAPAGRLCLNWLDAQSLASAPVSGAGNHSYCRNPDEDPRGWCYV 844

```

RESULT 2
US-60-197-873-13839
# Sequence 13839, Application US/60197873
# GENERAL INFORMATION:
# APPLICANT: Benjamin, Stephane
# APPLICANT: Tanaka, Hiroaki
# APPLICANT: Dumas Milne Edwards, Jean Baptiste
# APPLICANT: Joerbert, Severin
# APPLICANT: Giordano, Jean-Yves
# TITLE OF INVENTION: ESTs and Encoded Human Proteins
# FILE REFERENCE: 81 US1 PRO
# CURRENT APPLICATION NUMBER: US/60/197, 873
# CURRENT FILING DATE: 2000-04-18
# NUMBER OF SEQ ID NOS: 52153
# SOFTWARE: Patent.pm
# SEQ ID NO 13839
# LENGTH: 109
# TYPE: PRT
# ORGANISM: Homo sapiens
# FEATURE:
# NAME/KEY: SIGNAL
# LOCATION: -21...-1
# US-60-197-873-13839

```

Query Match	100.0%;	Score 357;	DB 27;	Length 109;
Best Local Similarity	100.0%;	Pred. No. 8.1e-34;		
Matches	60; Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;

QY 1 CFWNGHLYREDQTSPPAGRLCLNWLDAQGLASAPVSGAGNHSYCRNPDEDPRGWCYV 60
DQ 1 CFWNGHLYREDQTSPPAGRLCLNWLDAQGLASAPVSGAGNHSYCRNPDEDPRGWCYV 60
Dp 25 CFWNGHLYREDQTSPPAGRLCLNWLDAQGLASAPVSGAGNHSYCRNPDEDPRGWCYV 84

```

RESULT 3
PCT-US01-14827-8897
; Sequence 8897, Application PC/TUS0114827
; GENERAL INFORMATION:
; APPLICANT: HySeq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-104
; CURRENT APPLICATION NUMBER: PCT/US01/14827
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 16102
; SOFTWARE: Custom
; SEQ ID NO 8897
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (74)..(95)
; OTHER INFORMATION: KRINGLE DOMAIN SIGNATURE domain identified by EMATRIX.
; OTHER INFORMATION: accession number PR00018C, p-value=6.447e-17, raw score of 14.30
; NAME/KEY: DOMAIN
; LOCATION: (36)..(113)
; OTHER INFORMATION: kringle domain identified by Pfam, accession name kringle, E-
; OTHER INFORMATION: value=1.2e-05, Pfam score of 12.2
; CTT-US01-14827-8897

```

```
Query Match          100.0%; Score 357; DB 1; Length 117;  
Rest Local Similarity 100.0%; Pred.No.8,-8e-34;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
1 CFWDNHLYREDTSPAPGLRCLNNLNDOSGLASPVPVSGAGNHSYCNPNDEPRGWCYV 60  
|||||
```

Db 36 CFWDNGHLYREDQTSAPGLRCLNWLDAQSGLASAPVSGAGNHSCRNPDDEDPGPMCYV 95

```

RESULT 4
US-09-307-140-979
; Sequence 979, Application US/09307140
GENERAL INFORMATION:
APPLICANT: Geating, David P.
APPLICANT: Holtzman, Douglas A.
APPLICANT: Robison, Keith E.
TITLE OF INVENTION: Nucleic Acid Molecules Derived from a
TITLE OF INVENTION: Human Adipic Endothelium Library
FILE REFERENCE: MN98-14PA
CURRENT APPLICATION NUMBER: US/09/307,140
CURRENT FILING DATE: 1999-05-07
EARLIER APPLICATION NUMBER: 60/084,565
EARLIER FILING DATE: 1998-05-07
NUMBER OF SEQ ID NOS: 1168
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 979
LENGTH: 126
TYPE: prt
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)...(21)
US-09-307-140-979

```

Query Match	100.0%;	Score 357;	DB 17;	Length 126;
Best Local Similarity	100.0%;	Pred. No. 9.5e-34;		
Matches 60; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY 1 CFMONGHLYREDQTSAPGLRCLNWLDAOSGLASAPVSGAGNHSYCNPPDEDPGPMCYV 60
DB 25 CFMONGHLYREDQTSAPGLRCLNWLDAOSGLASAPVSGAGNHSYCNPPDEDPGPMCYV 84

```

RESULT 5
US-09-817-076-979
? Sequence 979, Application US/09817076
? GENERAL INFORMATION:
? APPLICANT: Gearling, David P.
? APPLICANT: Holtzman, Douglas A.
? APPLICANT: Robison, Keith E.
? TITLE OF INVENTION: Nucleic Acid Molecules Derived from a
? TITLE OF INVENTION: Human Aortic Endothelium Library
? FILE REFERENCE: MN99-189A
? CURRENT FILING DATE: 2001-03-26
? PRIOR APPLICATION NUMBER: US/09/817, 076
? PRIOR FILING DATE: 1999-05-07
? PRIOR APPLICATION NUMBER: 60/084,565
? PRIOR FILING DATE: 1998-05-07
? NUMBER OF SEQ ID NOS: 1168
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 979
? LENGTH: 126
? TYPE: prt
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: SIGNAL
? LOCATION: (1)...(21)
? US-09-817-076-979

```

Query Match	100.0%	Score 357	DB 22	Length 126
Best Local Similarity	100.0%	Pred. 9.5e-34		
Matches 60	Conservative 0	Mismatches 0	Indels 0	Gaps 0

CY	1	CPMDNGHLYREDQTSAPGRLCLNMLDQSLASAPVSGAGNHSYCRNPDEDFRGWCYV	60
cb	25	CPMDNGHLYREDQTSAPGRLCLNMLDQSLASAPVSGAGNHSYCRNPDEDFRGWCYV	84

```
RESULT 6 -
PCT-US02-05301-237
; Sequence 237, Application PC/TUS0205301
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: PS736PCT
; CURRENT APPLICATION NUMBER: PCT/US02/05301
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/304,417
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US 60/270,625
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 237
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (132)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (137)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (142)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (144)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US02-05301-237

Query Match          100.0%; Score 357; DB 1; Length 146;
Best Local Similarity 100.0%; Pred. No. 1,1e-33;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFWDNGHLYREDQTSPPAGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPFGPCV 60
Db 25 CFWDNGHLYREDQTSPPAGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPFGPCV 84

RESULT 7
PCT-US02-05301-311
; Sequence 311, Application PC/TUS0205301
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: PS736PCT
; CURRENT APPLICATION NUMBER: PCT/US02/05301
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/304,417
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US 60/270,625
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 311
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (132)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (137)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (142)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
PCT-US02-05301-311
```

```
; LOCATION: (144)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US02-05301-311

Query Match          100.0%; Score 357; DB 1; Length 146;
Best Local Similarity 100.0%; Pred. No. 1,1e-33;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFWDNGHLYREDQTSPPAGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPFGPCV 60
Db 25 CFWDNGHLYREDQTSPPAGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPFGPCV 84

RESULT 8
US-60-213-800-237
; Sequence 237, Application US/60213800
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CLO00707
; CURRENT APPLICATION NUMBER: US/60/213,800
; CURRENT FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 657
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 237
; LENGTH: 178
; TYPE: PRT
; ORGANISM: HUMAN
US-60-213-800-237

Query Match          100.0%; Score 357; DB 27; Length 178;
Best Local Similarity 100.0%; Pred. No. 1,4e-33;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFWDNGHLYREDQTSPPAGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPFGPCV 60
Db 25 CFWDNGHLYREDQTSPPAGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPFGPCV 84

RESULT 9
US-09-001-403-64
; Sequence 64, Application US/09001403
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Au-Young, Janice
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yue, Henry
; APPLICANT: Shah, Puri
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY PROTEINS
; NUMBER OF SEQUENCES: 150
; CORRESPONDENCE ADDRESS:
; ADDRESSES: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/001,403
; FILING DATE: HERREWITH
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
```

```

; NAME: BILLINGS, LUCY J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0455 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: LINDNOT05
; CLONE: 3122252
; US-09-001-403-64

Query Match          100.0%; Score 357; DB 1; Length 250;
Best Local Similarity 100.0%; Pred. No. 2e-33;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFWDNGHLYREDQTSAPGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPGPGWCYV 60
12 CFWDNGHLYREDQTSAPGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPGPGWCYV 71

RESULT 10
PCT-US00-27239-2
; Sequence 2, Application PC/TUS0027239
; GENERAL INFORMATION:
; APPLICANT: HUMAN GENOME SCIENCES, INC.
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
; FILE REFERENCE: PF378PCT2
; CURRENT APPLICATION NUMBER: PCT/US00/27239
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 09/411,977
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US00-27239-2

Query Match          100.0%; Score 357; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 2.1e-33;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFWDNGHLYREDQTSAPGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPGPGWCYV 60
25 CFWDNGHLYREDQTSAPGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPGPGWCYV 84

RESULT 11
PCT-US02-05301-164
; Sequence 164, Application PC/TUS0205301
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: PS736PCT
; CURRENT APPLICATION NUMBER: PCT/US02/05301
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/304,417
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US 60/270,625
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 164
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```

PCT-US02-05301-164

Query Match          100.0%; Score 357; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 2.1e-33;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFWDNGHLYREDQTSAPGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPGPGWCYV 60
25 CFWDNGHLYREDQTSAPGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPGPGWCYV 84

RESULT 12
PCT-US98-10728-2
; Sequence 2, Application PC/TUS9810728
; GENERAL INFORMATION:
; APPLICANT: EBNER, REINHARD
; APPLICANT: MOORE, PAUL
; APPLICANT: RUBEN, STEVEN
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/10728
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A.
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF379PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US98-10728-2

Query Match          100.0%; Score 357; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 2.1e-33;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFWDNGHLYREDQTSAPGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPGPGWCYV 60
25 CFWDNGHLYREDQTSAPGLRCLNWLDAQSGLASAPVSGAGNHSYCRNPDEDPGPGWCYV 84

RESULT 13
US-09-084-491A-2
; Sequence 2, Application US/09084491A
; GENERAL INFORMATION:
; APPLICANT: MOORE, PAUL A.
; APPLICANT: RUBEN, STEVEN M.
; APPLICANT: EBNER, REINHARD
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
```

STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/084,491A
FILING DATE: 27-MAY-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.
REGISTRATION NUMBER: 36,373
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8439
FAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-084-491A-2

Query Match 100.0%; Score 357; DB 14; Length 263;
Best Local Similarity 100.0%; Pred. No. 2.1e-33;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFMDNGHLYREDQTSAPGRCINMLDQSGLASAPVSGAGNHSYCRNPDDPRGPGWCV 60
DB 25 CFMDNGHLYREDQTSAPGRCINMLDQSGLASAPVSGAGNHSYCRNPDDPRGPGWCV 84

RESULT 14

US-09-144-889-2
Sequence 2, Application US/09144889B
GENERAL INFORMATION:
APPLICANT: Earl P. Albone
APPLICANT: Kristine K. Kirly
TITLE OF INVENTION: KRINGLE1
FILE REFERENCE: GH-70249
CURRENT APPLICATION NUMBER: US/09/144,889B
CURRENT FILING DATE: 1998-09-01
EARLIER APPLICATION NUMBER: 60/056,032
EARLIER FILING DATE: 1997-09-02
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 263
TYPE: PRT
ORGANISM: HOMO SAPIENS
US-09-144-889-2

Query Match 100.0%; Score 357; DB 15; Length 263;
Best Local Similarity 100.0%; Pred. No. 2.1e-33;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFMDNGHLYREDQTSAPGRCINMLDQSGLASAPVSGAGNHSYCRNPDDPRGPGWCV 60
DB 25 CFMDNGHLYREDQTSAPGRCINMLDQSGLASAPVSGAGNHSYCRNPDDPRGPGWCV 84

RESULT 15

US-09-611-526-3727
Sequence 3727, Application US/09611526
GENERAL INFORMATION:
APPLICANT: OTA, TOSHIO
APPLICANT: NISHIKAWA, TETSUO
APPLICANT: ISOGAI, TAKAO
APPLICANT: HAYASHI, KOJI

APPLICANT: ISHII, SHIZUKO
APPLICANT: KAWAI, YUKI
APPLICANT: WAKAMATSU, AI
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: NAGAI, KEIICHI
APPLICANT: KOJIMA, SHINICHI
APPLICANT: OTSUKI, TETSUJI
APPLICANT: KOGA, HISASHI
TITLE OF INVENTION: PRIMERS FOR SYNTHESIS OF FULL LENGTH CDNAS
TITLE OF INVENTION: AND THEIR USES
FILE REFERENCE: 08335/0122
CURRENT APPLICATION NUMBER: US/09/611,526
CURRENT FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: JP 1999-194486
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: JP 2000-118774
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: JP 2000-183765
PRIOR FILING DATE: 2000-05-02
NUMBER OF SEQ ID NOS: 4484
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3727
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
US-09-611-526-3727

Query Match 100.0%; Score 357; DB 20; Length 263;
Best Local Similarity 100.0%; Pred. No. 2.1e-33;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFMDNGHLYREDQTSAPGRCINMLDQSGLASAPVSGAGNHSYCRNPDDPRGPGWCV 60
DB 25 CFMDNGHLYREDQTSAPGRCINMLDQSGLASAPVSGAGNHSYCRNPDDPRGPGWCV 84

Search completed: April 7, 2003, 09:29:27
Job time : 52.2682 secs



GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:18:38 ; Search time 8.98129 Seconds
(without alignments)
1010.317 Million cell updates/sec

Title: US-10-057-951-2_COPY_25_84

Sequence: 1 CFWDNGHLYREDQTSPPAGL.....GNHSYCRNPDEDPGRCVCYV 60

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 679521 seqs, 151232488 residues

Total number of hits satisfying chosen parameters: 679521

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

1: Pending Patents AA New:*
2: /cgn2_6/pcdata/1/paa/US06_NEM_COMB.pep:*
3: /cgn2_6/pcdata/1/paa/US07_NEM_COMB.pep:*
4: /cgn2_6/pcdata/1/paa/US08_NEM_COMB.pep:*
5: /cgn2_6/pcdata/1/paa/US09_NEM_COMB.pep:*
6: /cgn2_6/pcdata/1/paa/US10_NEM_COMB.pep:*
7: /cgn2_6/pcdata/1/paa/US60_NEM_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	351	98.3	263	6	US-10-218-140-6002
2	280	78.4	264	6	US-10-144-779-488
3	146	40.9	293	7	US-60-452-680-14406
4	146	40.9	293	7	US-60-453-135-8958
5	146	40.9	293	7	US-60-450-8958
6	146	40.9	516	7	US-60-452-680-14405
7	146	40.9	516	7	US-60-453-135-8957
8	146	40.9	516	7	US-60-453-050-8957
9	146	40.9	527	5	US-09-612-314A-51
10	146	40.9	562	1	PCT-US02-32263-26
11	146	40.9	562	5	US-09-703-695A-4
12	146	40.9	562	6	US-10-287-994-26
13	146	40.9	562	7	US-60-452-680-14408
14	146	40.9	562	7	US-60-453-135-8960
15	146	40.9	562	7	US-60-453-050-8960
16	146	40.9	567	5	US-09-949-016-11501
17	141	39.5	655	1	PCT-US02-19017-28
18	141	39.5	656	7	US-60-453-135-11516
19	141	39.5	656	7	US-60-453-050-11516
20	135	37.8	482	7	US-60-452-680-14407
21	135	37.8	482	7	US-60-453-135-8959
22	135	37.8	482	7	US-60-453-050-8959
23	132	37.0	86	1	PCT-US02-27855-1
24	132	37.0	86	1	PCT-US02-27855A-1
25	132	37.0	86	6	US-10-233-675A-1
26	132	37.0	87	1	PCT-US02-27855-10

27	132	37.0	87	1	PCT-US02-27885A-10	Sequence 10, App1
28	132	37.0	87	6	US-10-233-675A-10	Sequence 10, App1
29	132	37.0	322	1	PCT-US02-27855-20	Sequence 20, App1
30	132	37.0	322	1	PCT-US02-27855-21	Sequence 21, App1
31	132	37.0	322	1	PCT-US02-27885A-20	Sequence 20, App1
32	132	37.0	322	1	PCT-US02-27885A-21	Sequence 21, App1
33	132	37.0	322	6	US-10-233-675A-20	Sequence 20, App1
34	132	37.0	322	6	US-10-233-675A-21	Sequence 21, App1
35	132	37.0	672	1	PCT-US02-27855-15	Sequence 15, App1
36	132	37.0	672	1	PCT-US02-27885A-15	Sequence 15, App1
37	132	37.0	674	1	US-10-233-675A-15	Sequence 15, App1
38	132	37.0	674	1	PCT-US02-27855-14	Sequence 14, App1
39	132	37.0	674	1	PCT-US02-27885A-14	Sequence 14, App1
40	132	37.0	674	6	US-10-233-675A-14	Sequence 14, App1
41	132	37.0	687	1	PCT-US02-27855-17	Sequence 17, App1
42	132	37.0	687	1	PCT-US02-27885A-17	Sequence 17, App1
43	132	37.0	687	6	US-10-233-675A-17	Sequence 17, App1
44	132	37.0	688	1	PCT-US02-27855-18	Sequence 18, App1
45	132	37.0	688	1	PCT-US02-27885A-18	Sequence 18, App1

ALIGNMENTS

```
RESULT 1
US-10-218-140-6002
; Sequence 6002, Application US/10218140
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYPEPTIDES
; FILE REFERENCE: 15966-543 CON
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 09/540,763
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/127,728
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/127,636
; PRIOR FILING DATE: 1999-04-02
; PRIOR APPLICATION NUMBER: 60/127,607
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 6322
; SOFTWARE: Curator Version 1.0
; SEQ ID NO 6002
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-218-140-6002
Query Match          98.3%; Score 351; DB 6; Length 263;
Best Local Similarity 98.3%; Pred. No. 4.5e-29;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CFWDNGHLYREDQTSPPAGLRCINWLDAGSLAPVSGAGNHSYCRNPDEDPGRCVCYV 60
Db 25 CFWDNGHLYREDQTSPPAGLRCINWLDAGSLAPVSGAGNHSYCRNPDEDPGRCVCYV 84
RESULT 2
US-10-144-779-488
; Sequence 488, Application US/10144779
; GENERAL INFORMATION:
; APPLICANT: SUBRAMANIAN, Mani et al.
; TITLE OF INVENTION: MOUSE ORTHOLOGS OF HUMAN DISEASE GENES,
; FILE REFERENCE: C1001235
; CURRENT APPLICATION NUMBER: US/10/144,779
; CURRENT FILING DATE: 2002-05-15
; NUMBER OF SEQ ID NOS: 864
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 488
```

LENGTH: 264
TYPE: PRT
ORGANISM: Mus musculus
US-10-144-779-488

Query Match 78.4%; Score 280; DB 6; Length 264;
Best Local Similarity 75.0%; Pred. No. 1.3e-21;
Matches 45; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 1 CFWDNGHLYREDQTSPPAGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 60
DB 25 CFWDNGHLYREDQTSPPAGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 84

RESULT 3
US-60-452-680-14406

Sequence 14406, Application US/60452680
GENERAL INFORMATION:

APPLICANT: CARGILL, Michele

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

FILE REFERENCE: CLO01450

CURRENT APPLICATION NUMBER: US/60/452,680

NUMBER OF SEQ ID NOS: 116213

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 14406

LENGTH: 293

TYPE: PRT

ORGANISM: Homo sapiens

US-60-452-680-14406

Query Match 40.9%; Score 146; DB 7; Length 293;
Best Local Similarity 43.5%; Pred. No. 1.6e-07;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

QY 1 CFWDNGHLYREDQTSPPAGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 51
DB 127 CYEDGISTYRGWTSTASGAECTNW--NSSALAQKPYSGRRPDATRLGIGNHNYCRNPDR 184

QY 52 DPRGWCYV 60
DB 185 DSK-PWCYV 192

RESULT 4
US-60-453-135-8958

Sequence 8958, Application US/60453135
GENERAL INFORMATION:

APPLICANT: CARGILL, Michele

APPLICANT: IAKOUBOVA, Olga

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

FILE REFERENCE: CLO01456

CURRENT APPLICATION NUMBER: US/60/453,135

NUMBER OF SEQ ID NOS: 82762

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 8958

LENGTH: 293

TYPE: PRT

ORGANISM: Homo sapiens

US-60-453-135-8958

Query Match 40.9%; Score 146; DB 7; Length 293;
Best Local Similarity 43.5%; Pred. No. 1.6e-07;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

QY 1 CFWDNGHLYREDQTSPPAGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 51
DB 127 CYEDGISTYRGWTSTASGAECTNW--NSSALAQKPYSGRRPDATRLGIGNHNYCRNPDR 184

QY 52 DPRGWCYV 60
DB 185 DSK-PWCYV 192

RESULT 5
US-60-453-050-8958

Sequence 8958, Application US/60453050
GENERAL INFORMATION:

APPLICANT: CARGILL, Michele

APPLICANT: LUKE, May

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

FILE REFERENCE: CLO01457

CURRENT APPLICATION NUMBER: US/60/453,050

NUMBER OF SEQ ID NOS: 82762

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 8958

LENGTH: 293

TYPE: PRT

ORGANISM: Homo sapiens

US-60-453-050-8958

Query Match 40.9%; Score 146; DB 7; Length 293;
Best Local Similarity 43.5%; Pred. No. 1.6e-07;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

QY 1 CFWDNGHLYREDQTSPPAGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 51
DB 127 CYEDGISTYRGWTSTASGAECTNW--NSSALAQKPYSGRRPDATRLGIGNHNYCRNPDR 184

QY 52 DPRGWCYV 60
DB 185 DSK-PWCYV 192

RESULT 6
US-60-452-680-14405

Sequence 14405, Application US/60452680
GENERAL INFORMATION:

APPLICANT: CARGILL, Michele

APPLICANT: GRUPE, Andrew

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

FILE REFERENCE: CLO01450

CURRENT APPLICATION NUMBER: US/60/452,680

NUMBER OF SEQ ID NOS: 116213

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 14405

LENGTH: 516

TYPE: PRT

ORGANISM: Homo sapiens

US-60-452-680-14405

Query Match 40.9%; Score 146; DB 7; Length 516;
Best Local Similarity 43.5%; Pred. No. 2.6e-07;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

QY 1 CFWDNGHLYREDQTSPPAGRLCLNWLDAOSGLASAPVS-----GAGNHSYCRNPDE 51
DB 81 CYEDGISTYRGWTSTASGAECTNW--NSSALAQKPYSGRRPDATRLGIGNHNYCRNPDR 138

QY 52 DPRGWCYV 60
DB 139 DSK-PWCYV 146

RESULT 7
US-60-453-135-8957
Sequence 8957, Application US/60453135
GENERAL INFORMATION:

```

? APPLICANT: CARGILL, Michele
? APPLICANT: IAKOUBOVA, Olga
? TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
? TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
? FILE REFERENCE: CL001456
? CURRENT APPLICATION NUMBER: US/60/453,135
? CURRENT FILING DATE: 2003-03-10
? NUMBER OF SEQ ID NOS: 82762
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 8957
? LENGTH: 516
? TYPE: PRT
? ORGANISM: Homo sapiens
? US-60-453-135-8957

```

Query Match	40.9%;	Score 146;	DB 7;	Length 516;
Best Local Similarity	43.5%;	Pred. No. 2.6e-07;		
Matches 30;	Conservative 5;	Mismatches 22;	Indels 12;	Gaps 3;

DQ 1 CFPMNGHLVREDQTSAPGLRCLNWLDAQSGLASAPVS-----GAGHSYCRAPDE 51
| : | | | : | | | | | | | | | |
D8 81 CYEDQGISIRGTWSTAESGAECTNW--NSSALAQKPYSGRPDAIRLGLGNHNYCRAPDR 138

QY	52	DPRGPWCYV	60
		:	
Db	139	DSK-PWCYV	146

RESULT 8
US-60-453-050-8957

Sequence 8957, Application U
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
APPLICANT: LUKE, May
END OF INFORMATION.

```

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001457
; CURRENT APPLICATION NUMBER: US/60/453, 050
; CURRENT FILING DATE: 2003-03-10

```

```

; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8957
; LENGTH: 516

```

ORGANISM: Homo sapiens
US-60-453-050-8957

Query Match	40.9%;	Score 146;	DB 7;	Length 516;
Best Local Similarity	43.5%;	Pred. No. 2.6e-07;		
Matches 30; Conservative	5;	Mismatches 22;	Indels 12;	Gaps 3

```

QY      1 CFFDNGHLTYREDQTSAPGLKRLNWEDAQSGLSAPVS-----GAGNHSYCRNDE 51
      | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      81 CYEDDGISYRGTSWAESGAECTNW--NSSALAQKPYSGRRDPAIRLGLGNINYCRNDR 138

```

QY	52	DPRGPWCYV	60
		:	
Db	139	DSK-PWCYV	146

RESULT 9
US-09-612-314A-51

APPLICANT: SMITH, RICHARD ANTHONY GODWIN
APPLICANT: DODD, IAN

; APPLICANT: MOSSAKOWSKA, DANUTA EVA IRENA
 ; TITLE OF INVENTION: CONJUGATES OF SOLUBLE PEPTIDIC COMPOUNDS WITH
 ; TITLE OF INVENTION: MEMBRANE-BINDING AGENTS
 ; FILE REFERENCE: 37945-0004

```

; CURRENT APPLICATION NUMBER: US/09/612,314A
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 09/214,913

```

```

:
: PRIORITY FILING DATE: 1999-03-16
: PRIOR APPLICATION NUMBER: PCT/EP97/03715
: PRIOR FILING DATE: 1997-07-08
: PRIOR APPLICATION NUMBER: GB 96 148 71.3
: PRIOR FILING DATE: 1996-07-15
: NUMBER OF SEQ ID NOS: 53
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 51
: LENGTH: 527
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: tissue plasminogen activator
US-09-612-314A-51

```

Query Match	40.9%	Score 146;	DB 5;	Length 527;
Best Local Similarity	43.5%	Pred. No. 2.6e-07;		
Matches 30; Conservative	5;	Mismatches 22;	Indels 12;	Gaps 3;

D5
92 CYEDQGISYRGTSVAESGAECTNW--NSSALAQKPYSGRRPDATRLGLGNHNYCRNPDR 149

D6
92 CFVNDNGHLRYREDQTSPPAGELCLMTLDAQSGLASAPVS-----GAGHSYCRNPDE 51

QY	52	DPRGFWCYV	60
		:	
Db	150	DSK-FWCYV	157

RESULT 10
PCT-US02-32263-26

```

; Sequence 26, Application PC/TUS023222
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc
; APPLICANT: Defrees, Shawn

```

```
;; APPLICANT: Zopf, David
;; APPLICANT: Bayer, Robert
;; APPLICANT: Bowe, Caryn
;; APPLICANT: Hakes, David
;; APPLICANT: Chen, Xi
```

TITLE OF INVENTION: REMODELING AND GLYCOCONJUGATION OF PEPTIDES
 FILE REFERENCE: 040853-01-5050WO
 CURRENT APPLICATION NUMBER: PCT/US02/32263
 CURRENT FILING DATE: 2002-10-31

PRIOR APPLICATION NUMBER: US 60/328,523
PRIOR FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: US 60/344,692
PRIOR FILING DATE: 2001-10-19

PRIOR APPLICATION NUMBER: US 60/334,233
PRIOR FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: US 60/334,302
PRIOR FILING DATE: 2001-11-08

PRIOR FILING DATE: 2002-11-06
PRIOR APPLICATION NUMBER: US 60/387,292
PRIOR FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: US 60/391,777
PRIOR FILING DATE: 2002-06-25

PRIOR FILING DATE: 2002-08-23
PRIOR APPLICATION NUMBER: US 60/396,594
PRIOR FILING DATE: 2002-07-17
PRIOR APPLICATION NUMBER: US 60/404,249
PRIOR FILING DATE: 2002-08-16

```

;
; PRIORITY FILING DATE: 2002-08-16
;
; PRIOR APPLICATION NUMBER: US 60/407,527
;
; PRIOR FILING DATE: 2002-08-28
;
; NUMBER OF SEQ ID NOS: 62
;
; SOFTWARE: DataBlast version 3.1

```

```

; SOFTWARE: PatentInversion 3.1
; SEQ ID NO 26
; LENGTH: 562
; TYPE: PRT
; ORIGIN: Homo sapiens

```

PCT-US02-32263-26	
; ORGANISM: Homo sapiens	
Query Match	40.9%
Score 146; DB 1; Length 562;	

Best Local Similarity	43.5%;	Pred. No. 2.8e-07;
Matches	30;	Conservative
	5;	Mismatches
	22;	Indels
	12;	Gaps
	3;	

[illegible]

Mon Apr 7 10:22:30 2003

us-10-057-951-2_copy_25_84.rapn

Page 5

```
Db 127 CYEDOGISYRGWSTAESGAECTNW--NSSALAQKPYSGRRPDATRLGIGNHNYCRNPDR 184
QY 52 DPERGPMCTV 60
| : |||
Db 185 DSK-PMCTV 192
```

```
RESULT 15
US-60-453-050-8960
; Sequence 8960, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8960
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-8960
```

```
Query Match 40.9%; Score 146; DB 7; Length 562;
Best Local Similarity 43.5%; Pred. No. 2.8e-07;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;
```

```
QY 1 CFMNGHLYREDQSPARGLRGLAWLDAGSLAPVS-----GAGNHSYCRNPDE 51
Db 127 CYEDOGISYRGWSTAESGAECTNW--NSSALAQKPYSGRRPDATRLGIGNHNYCRNPDR 184
QY 52 DPERGPMCTV 60
| : |||
Db 185 DSK-PMCTV 192
```

Search completed: April 7, 2003, 09:30:48
Job time : 9.98129 secs

10

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:28 ; Search time 5.98753 Seconds
(without alignments)
294.842 Million cell updates/sec

Title: US-10-057-951-2_COPY_25_84

Perfect score: 357
Sequence: 1 CFWDNGHLYREDQTSFAPGL.....GNHSTCRNPDDEPRGWCYV 60

Scoring table:
BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
4: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/iaa/Dackfil1est.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	357	100.0	263	4 US-09-411-977-2	Sequence 2, Appl
2	150	42.0	472	2 US-08-811-949-63	Sequence 63, Appl
3	146	40.9	83	2 US-08-811-949-2	Sequence 2, Appl
4	146	40.9	437	2 US-08-811-949-49	Sequence 49, Appl
5	146	40.9	437	2 US-08-811-949-51	Sequence 51, Appl
6	146	40.9	437	2 US-08-811-949-55	Sequence 55, Appl
7	146	40.9	437	2 US-08-811-949-57	Sequence 57, Appl
8	146	40.9	527	1 US-07-609-510B-16	Sequence 16, Appl
9	146	40.9	527	2 US-08-811-949-39	Sequence 39, Appl
10	146	40.9	527	5 PCT-US91-01025A-2	Sequence 2, Appl
11	146	40.9	527	5 5185259-8	Patent No. 5185259
12	146	40.9	527	6 520913-1	Patent No. 520913
13	146	40.9	546	6 5200340-6	Patent No. 5200340
14	146	40.9	562	2 US-08-811-949-43	Sequence 43, Appl
15	146	40.9	562	2 US-08-560-098A-50	Sequence 50, Appl
16	146	40.9	562	2 US-08-883-795A-38	Sequence 38, Appl
17	146	40.9	562	6 5185259-3	Patent No. 5185259
18	146	40.9	562	6 5200340-2	Patent No. 5200340
19	146	40.9	562	6 5344773-2	Patent No. 5344773
20	141	39.5	655	1 US-08-148-910-12	Sequence 12, Appl
21	141	39.5	655	1 US-08-448-937A-12	Sequence 12, Appl
22	139	38.5	356	1 US-08-427-640-8	Sequence 8, Appl
23	134	37.5	477	2 US-08-560-098A-51	Sequence 51, Appl
24	131	36.7	138	2 US-08-797-689-12	Sequence 12, Appl
25	131	36.7	365	1 US-08-093-741-83	Sequence 83, Appl
26	131	36.7	365	1 US-08-720-012-83	Sequence 83, Appl
27	131	36.7	393	2 US-08-560-098A-44	Sequence 44, Appl

28	131	36.7	393	4 US-08-967-024C-24	Sequence 24, Appl
29	131	36.7	393	4 US-08-967-024C-25	Sequence 25, Appl
30	131	36.7	411	1 US-08-087-163-1	Sequence 1, Appl
31	131	36.7	411	1 US-08-286-748B-18	Sequence 18, Appl
32	131	36.7	411	1 US-08-153-799-18	Sequence 18, Appl
33	131	36.7	411	2 US-08-560-098A-48	Sequence 48, Appl
34	131	36.7	411	4 US-09-181-816-1	Sequence 1, Appl
35	131	36.7	430	1 US-07-942-157A-3	Sequence 3, Appl
36	131	36.7	430	6 5219569-2	Patent No. 5219569
37	131	36.7	431	6 5188829-1	Patent No. 5188829
38	131	36.7	432	2 US-08-560-098A-47	Sequence 47, Appl
39	128	35.9	355	1 US-08-427-640-6	Sequence 6, Appl
40	127	35.6	326	4 US-09-411-977-3	Sequence 3, Appl
41	127	35.6	347	2 US-08-811-949-1	Sequence 1, Appl
42	127	35.6	354	2 US-08-811-949-61	Sequence 61, Appl
43	127	35.6	355	1 US-08-137-116-1	Sequence 1, Appl
44	127	35.6	355	1 US-08-217-618-1	Sequence 1, Appl
45	127	35.6	355	1 US-08-427-640-2	Sequence 2, Appl

ALIGNMENTS

```
RESULT 1
US-09-411-977-2
; Sequence 2, Application US/09411977
; Patent No. 6372473
; GENERAL INFORMATION:
; APPLICANT: Moore, Paul A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
; FILE REFERENCE: PR378P1
; CURRENT APPLICATION NUMBER: US/09/411,977
; CURRENT FILING DATE: 1998-10-04
; EARLIER APPLICATION NUMBER: 09/084,491
; EARLIER FILING DATE: 1998-05-27
; EARLIER APPLICATION NUMBER: 60/048,000
; EARLIER FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-411-977-2

Query Match      100.0%; Score 357; DB 4; Length 263;
Best Local Similarity 100.0%; Pred. No. 9.6e-36;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CFWDNGHLYREDQTSFAPGLRCLNWLDAQSLASAPVSGAGNHSYCRNPDDEPRGWCYV 60
Db      25 CFWDNGHLYREDQTSFAPGLRCLNWLDAQSLASAPVSGAGNHSYCRNPDDEPRGWCYV 84

RESULT 2
US-08-811-949-63
; Sequence 63, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIMA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTARI, JOUIT
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, WAIR & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
```

```

CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ. ID NO.: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 472 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-63

```

```

Query Marc3 42.0%; Score 150; DB 2; Length 472;
Best Local Similarity 44.9%; Pred No 2,5e-10;
Matches 31; Conservative 4; Mismatches 22; Indels 12; Gaps 3

Qy 1 CPMDNGLHYREDOTSPAPGRLCLNWDIAQSGIASAPVS-----GAGNHSYCRNPDE 51
      | : . | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 37 CYEODGISTYRGVLTWTAESGAECTNW--NSSALAQKPSGRRRDPDIRLGLGNHNYCRNPDR 94
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 52 DPGPMCYV 60
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 95 DSK-PWCYV 102

RESULT 3
US-08-811-949-2
; Sequence 2, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIVA, MINEO
; APPLICANT: SATTO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOHAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.

```

```

:      REGISTRATION NUMBER: 24_618
:      REFERENCE/DOCKET NUMBER: 18-966-0
:      TELECOMMUNICATION INFORMATION:
:      TELEPHONE: 703-413-3000
:      TELEFAX: 703-413-2220
:      INFORMATION FOR SEQ ID NO: 2:
:      SEQUENCE CHARACTERISTICS:
:      LENGTH: 83 amino acids
:      type: amino acid
:      STRANDEDNESS: single
:      TOPOLOGY: linear
:      MOLECULE TYPE: peptide
:      US-08-811-943-2

```

Query Match	40.9%;	Score 146;	DB 2;	Length 83;
Best Local Similarity	43.5%;	Pred. No. 9.8e-11;		
Matches	30;	Conservative	5;	Mismatches 22; Indels 12; Gaps 3
QY	1	CFMDNGHLYAEEDQTSFAPGLRCLNMLDQSLGASAPS-----GASHSYSCNPDE	51	
Db	1	CYEOGISTRGWTSIAESGAECTNM--NSSALQKPIFSGRRPDAIRLGLGNHNYCRNPR	58	
QY	52	DPRGPMCVY	60	
Db	59	DSK-PMCVY	66	

RESULT 4
US-08-811-949-49
; Sequence 49, Application US/08811949

GENERAL INFORMATION:
APPLICANT: NIWA, MINBO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HIITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR

1 CORRESPONDENCE ADDRESS:
2 ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
3 ADDRESSEE: P.C.
4 STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
5 CITY: ARLINGTON
6 STATE: VA
7 COUNTRY: USA
8 ZIP: 22202
9 COMPUTER READABLE FORM:
10 MEDIUM TYPE: floppy disk
11 COMPUTER: IBM PC compatible
12 OPERATING SYSTEM: PC-DOS/MS-DOS
13 SOFTWARE: PatentIn Release #1.0, Version #1.30
14 CURRENT APPLICATION DATA:
15 APPLICATION NUMBER: US/08/811,949
16 FILING DATE: 05-MAR-1997
17 CLASSIFICATION: 435
18 ATTORNEY/AGENT INFORMATION:
19 NAME: OBLON, NORMAN F.
20 REGISTRATION NUMBER: 24,618
21 REFERENCE/DOCKET NUMBER: 18-966-0
22 TELECOMMUNICATION INFORMATION:
23 TELEPHONE: 703-413-3000
24 TELEFAX: 703-413-2220
25 INFORMATION FOR SEQ ID NO: 49:
26 SEQUENCE CHARACTERISTICS:
27 LENGTH: 437 amino acids
28 TYPE: amino acid
29 TOPOLOGY: linear
30 MOLECULE TYPE: protein
31 US-08-811-949-49

Query Match 40.9%; Score 146; DB 2; Length 437;

STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-57

Query Match 40.9%; Score 146; DB 2; Length 437;
Best Local Similarity 43.5%; Pred. No. 7.1e-10;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

QY 1 CFWDNGHLYREDQTSAPGRLCLNWLDAQSLASAPVS-----GAGNHSYCRNPDE 51
DB 2 CYBDQGISYRGWTSAISGAECTNM--NSSALAQKPYSGRRPDARLGLGNHNYCRNPDR 59

QY 52 DPRGPWCYV 60
DB 60 DSK-PWCYV 67

RESULT 8
US-07-609-510B-16
Sequence 16, Application US/07609510B
Patent No. 5326700
GENERAL INFORMATION:
APPLICANT: Berg et al.
TITLE OF INVENTION: Method for Altering Post-Translational Processing of Tissue Pl
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: IN.
COUNTRY: U.S.A.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/609,510B
FILING DATE: 19901106
CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-609-510B-16

Query Match 40.9%; Score 146; DB 1; Length 527;
Best Local Similarity 43.5%; Pred. No. 8.9e-10;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

QY 1 CFWDNGHLYREDQTSAPGRLCLNWLDAQSLASAPVS-----GAGNHSYCRNPDE 51
DB 92 CYBDQGISYRGWTSAISGAECTNM--NSSALAQKPYSGRRPDARLGLGNHNYCRNPDR 149

QY 52 DPRGPWCYV 60
DB 150 DSK-PWCYV 157

RESULT 9
US-08-811-949-39
Sequence 39, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NINAK, MINO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JODI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-39

Query Match 40.9%; Score 146; DB 2; Length 527;
Best Local Similarity 43.5%; Pred. No. 8.9e-10;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

QY 1 CFWDNGHLYREDQTSAPGRLCLNWLDAQSLASAPVS-----GAGNHSYCRNPDE 51
DB 92 CYBDQGISYRGWTSAISGAECTNM--NSSALAQKPYSGRRPDARLGLGNHNYCRNPDR 149

QY 52 DPRGPWCYV 60
DB 150 DSK-PWCYV 157

RESULT 10
PCT-US91-01025A-2

```
Sequence 2, Application PC/TUS9101025A
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: Tissue Plasminogen Activator Having Fibrin
; TITLE OF INVENTION: Specific Properties
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESSES:
; ADDRESS: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; City: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/01025A
; FILING DATE: 19910214
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/486,657
; FILING DATE: 1 March 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 454P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/265-1896
; TELEFAX: 415/265-9881
; TRFEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
PCT-US91-01025A-2

Query Match 40.9%; Score 146; DB 5; Length 527;
Best Local Similarity 43.5%; Pred. No. 8.9e-10;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

Qy 1 CFMNGHLVREDQTSAPAGRLCLNWLDAQSGLASAPVS-----GAGNSYCRNPDE 51
Db 92 CYEDQGISYRGWTSTAESGAECTNW--NSSALAQKPYSGRRPDAIRLGLGNHYCRNPDR 149
Qy 52 DPGPWCYV 60
Db 150 DSK-PWCYV 157

RESULT 11
5185259-8
; Patent No. 5185259
; APPLICANT: GOEDDEL, DAVID V.; KOHR, WILLIAM J.; PENNICA, DIANE;
; VEHAR, GORDON A.
; TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN
; ACTIVATOR
; NUMBER OF SEQUENCES: 15
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/489,855
; FILING DATE: 02-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 12,694
; FILING DATE: 09-FEB-1987
; APPLICATION NUMBER: 483,052
; FILING DATE: 07-APR-1983
; APPLICATION NUMBER: 398,003
; FILING DATE: 14-JUL-1982
; APPLICATION NUMBER: 374,860
; FILING DATE: 05-MAY-1982
```

```
; SEQ ID NO:8:
; LENGTH: 527
5185259-8
Query Match 40.9%; Score 146; DB 6; Length 527;
Best Local Similarity 43.5%; Pred. No. 8.9e-10;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

Qy 1 CFMNGHLVREDQTSAPAGRLCLNWLDAQSGLASAPVS-----GAGNSYCRNPDE 51
Db 92 CYEDQGISYRGWTSTAESGAECTNW--NSSALAQKPYSGRRPDAIRLGLGNHYCRNPDR 149
Qy 52 DPGPWCYV 60
Db 150 DSK-PWCYV 157

RESULT 12
5520913-1
; Patent No. 5520913
; APPLICANT: ANDERSON, STEPHEN; BENNETT, WILLIAM F.; BOTSTEIN,
; DAVID; HIGGINS, DEBORAH L.; PAONI, NICHOLAS F.; ZOLLER, MARK J.
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR HAVING
; ZYMOGENIC PROPERTIES
; NUMBER OF SEQUENCES: 35
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/88,451
; FILING DATE: 06-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 770,510
; FILING DATE: 03-OCT-1991
; APPLICATION NUMBER: 384,608
; FILING DATE: 24-JUL-1989
; APPLICATION NUMBER: 240,856
; FILING DATE: 02-SEP-1988
; SEQ ID NO:1:
; LENGTH: 527
5520913-1

Query Match 40.9%; Score 146; DB 6; Length 527;
Best Local Similarity 43.5%; Pred. No. 8.9e-10;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

Qy 1 CFMNGHLVREDQTSAPAGRLCLNWLDAQSGLASAPVS-----GAGNSYCRNPDE 51
Db 92 CYEDQGISYRGWTSTAESGAECTNW--NSSALAQKPYSGRRPDAIRLGLGNHYCRNPDR 149
Qy 52 DPGPWCYV 60
Db 150 DSK-PWCYV 157

RESULT 13
5200340-6
; Patent No. 5200340
; APPLICANT: FOSTER, DONALD C.; MULVHILL, EILEEN R.; O'HARA,
; PATRICK J.; PINDEL, KURT; YOSHITAKE, SHINJI
; TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
; ACTIVATORS
; NUMBER OF SEQUENCES: 34
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/53,412
; FILING DATE: 22-MAY-1987
; SEQ ID NO:6:
; LENGTH: 546
5200340-6

Query Match 40.9%; Score 146; DB 6; Length 546;
Best Local Similarity 43.5%; Pred. No. 9.2e-10;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;

Qy 1 CFMNGHLVREDQTSAPAGRLCLNWLDAQSGLASAPVS-----GAGNSYCRNPDE 51
Db 92 CYEDQGISYRGWTSTAESGAECTNW--NSSALAQKPYSGRRPDAIRLGLGNHYCRNPDR 149
```

Mon Apr 7 10:22:27 2003

us-10-057-951-2_copy_25_84.rai

Page 6

Db 127 CYEDQGISYRGWTSAESGAECTNW--NSSALAQKPYSGRRPAIRLIGNNHYCRNPDR 184
Qy 52 DPGPWCYV 60
Db 185 DSK-PMCYV 192

RESULT 14
US-08-811-949-43
Sequence 43, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIMA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESSES:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 562 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-08-811-949-43

Query Match 40.9%; Score 146; DB 2; Length 562;
Best Local Similarity 43.5%; Pred. No. 9.6e-10;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;
Qy 1 CFWDNGHLYREDQTSPPAGRLCLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 51
Db 127 CYEDQGISYRGWTSAESGAECTNW--NSSALAQKPYSGRRPAIRLIGNNHYCRNPDR 184
Qy 52 DPGPWCYV 60
Db 185 DSK-PMCYV 192

RESULT 15
US-08-560-098A-50
Sequence 50, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WENDEL, Stephan
APPLICANT: HEINZEL-WITLAND, Regina
APPLICANT: STEFFENS, Gerd Josef

TITLE OF INVENTION: Proteins having fibrinolytic and
TITLE OF INVENTION: Coagulation-inhibiting Properties
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Evensen, Mckeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELEPHONE: (202) 628-8844
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 562 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-08-560-098A-50

Query Match 40.9%; Score 146; DB 2; Length 562;
Best Local Similarity 43.5%; Pred. No. 9.6e-10;
Matches 30; Conservative 5; Mismatches 22; Indels 12; Gaps 3;
Qy 1 CFWDNGHLYREDQTSPPAGRLCLNWLDAQSGLASAPVS-----GAGNHSYCRNPDE 51
Db 127 CYEDQGISYRGWTSAESGAECTNW--NSSALAQKPYSGRRPAIRLIGNNHYCRNPDR 184
Qy 52 DPGPWCYV 60
Db 185 DSK-PMCYV 192

Search completed: April 7, 2003, 09:22:30
CPU time : 6.98753 secs

m/azpccimrncal source: BRIAN BIRLOI N2
 C;Genetics:

A:Gene: CESP.F45E12.2
A:Introns: 55/1; 152/3; 392/2; 650/3; 691/3; 731/3
C:Superfamily: transcription initiation factor IIIB 90K chain, transcription initiation

Query Match 9.6%; Score 87.5; DB 2; Length 759;
Best Local Similarity 30.4%; Pred. No. 4.2;
Matches 28; Conservative 17; Mismatches 32; Indels 15; Gaps 5;

QY 1 SGEAGVPEKPCEDL--RCPEPTTQALPAFTTEIQDASEGPADEV-----QVFAPNAL 53
DB 577 ASESTIQKRSIFDLTEECSEFSKNSSPKYNLKVESAS--PSTSEVSSIEHKFVP----- 630
QY 54 PARSEAAAVQPIGTSQRVNMSEKKDLGTL 85
DB 631 PARSRVAKVPIIGAKKLALN--EVKNVHTV 660

RESULT 3
A6510
leucyl tRNA synthetase [imported] - Chlamydia pneumoniae (strain J138)
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: A6510
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ie
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A65491; MUID:20330349; PMID:10871362
A:Accession: A6510
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-820 <STO>
A:Cross-references: GB:BA000008; NID:g8978526; PIDN:BA98363.1; GSPDB:GN00142
A:Experimental source: strain J138
C:Genetics:
A:Gene: leuS
C:Superfamily: leucine-tRNA ligase

Query Match 9.4%; Score 86.5; DB 2; Length 820;
Best Local Similarity 24.1%; Pred. No. 5.7;
Matches 38; Conservative 28; Mismatches 49; Indels 43; Gaps 9;

QY 20 TTGQALPAFTTEIQE-----ASEGPADEVQVFAPNALPARSEAA--VQPIGTS 69
DB 235 TQGSLEAFTTRLDTLGVSFLVIAPEHPDLSIV-----SEQRDEVTAIVQESLRKS 288
QY 70 QRVNMS-KEKDLGTLGY-----VLGITMV-----IIAIGAGIILGYSYKRGKDLKQ 119
DB 289 ERDRISSVKTKTGVTGNVAKHPITGNLIPWISDVVLGYGTGVVMGV-----PA 339
QY 120 HDQKVCEREMQRTLPISAFNPTCEIVDEKTVVHTS 157
DB 340 HDER--DREPAEM-----FSLPIHEVIDDNGVCISHN 369

RESULT 4
C72113
leucine-tRNA ligase (EC 6.1.1.4) [similarity] - Chlamydia pneumoniae (strains CWL029
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 03-Jun-2002
C:Accession: C72113; F81557
R:Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606; PMID:10192388
A:Accession: C72113
A:Molecule type: DNA
A:Residues: 1-820 <ARN>
A:Cross-references: GB:AE001602; GB:AE001363; NID:g4376416; PIDN:AA18306.1; PID:g437642
A:Experimental source: strain CWL029
R:Read, T.D.; Brinham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.

A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: F81557
A:Molecule type: DNA
A:Residues: 1-820 <REA>
A:Cross-references: GB:AE002219; GB:AE002161; NID:g7189524; PIDN:AAJ8433.1; PID:g718952
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: leuS; CP0618
C:Superfamily: leucine-tRNA ligase
C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 9.4%; Score 86.5; DB 2; Length 820;
Best Local Similarity 24.1%; Pred. No. 5.7;
Matches 38; Conservative 28; Mismatches 49; Indels 43; Gaps 9;

QY 20 TTGQALPAFTTEIQE-----ASEGPADEVQVFAPNALPARSEAA--VQPIGTS 69
DB 235 TQGSLEAFTTRLDTLGVSFLVIAPEHPDLSIV-----SEQRDEVTAIVQESLRKS 288
QY 70 QRVNMS-KEKDLGTLGY-----VLGITMV-----IIAIGAGIILGYSYKRGKDLKQ 119
DB 289 ERDRISSVKTKTGVTGNVAKHPITGNLIPWISDVVLGYGTGVVMGV-----PA 339
QY 120 HDQKVCEREMQRTLPISAFNPTCEIVDEKTVVHTS 157
DB 340 HDER--DREPAEM-----FSLPIHEVIDDNGVCISHN 369

RESULT 5
T31432
K-Cl cotransport protein 2, furosemide-sensitive - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 07-Dec-1999
C:Accession: T31432
R:Payne, U.A.; Stevenson, T.J.; Donaldson, L.F.
J. Biol. Chem. 271, 16245-16252, 1996
A:Title: Molecular characterization of a putative K-Cl cotransporter in rat brain: a new
A:Reference number: Z21031; MUID:96279171; PMID:8663311
A:Accession: T31432
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1116 <PAY>
A:Cross-references: EMBL:U55816; NID:g1403708; PID:g1403709; PIDN:AAC52635.1
A:Experimental source: strain Sprague Dawley; clone ERB10; 5ERB12; brain
C:Genetics:
A:Gene: KCC2
C:Keywords: transmembrane protein

Query Match 9.3%; Score 85.5; DB 2; Length 1116;
Best Local Similarity 27.1%; Pred. No. 10;
Matches 32; Conservative 15; Mismatches 44; Indels 27; Gaps 4;

QY 79 KKDGLTGLVYGLTMMVITIAIGAGIILGYSY-----KRGKDLKQGH--DQKVCEREMQ 130
DB 872 KKDGLTFLHLRTAEVEVEMHESDISAYTYEKTLMQESQILKQWHLTKEREHEIQ 931
QY 131 RIT-----LPLSAFTNPTC---EIVDEKTVVHTSQTVPDPOGSGTP 169
DB 932 SITDSRGSIRKKNPANRLNLNPEIRACDNEKPEEYQVLIHQDAPSPSSPSP 989

RESULT 6
E82525
primosomal protein N' XP2689 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: E82525
R:Anonymous. The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A>Note: For a complete list of authors see reference number A59328 below
A:Accession: E82525

Oy	31	EIOEASEPQADDEVVFPANALPARSEAAAVOPVIGISQRYMNSKEKDLGTL--GYV	88
Dd	227	EVTPTKEPPPAKPMTETXAKAKADTEENKA-----PSIGEQTIVRVDRRLDHMLNIGSLV	282
Oy	89	LGITMMVIIIAAGAILIGSYTKRGKD.LKEQHDKVCYCREMQRTLP.SATNPNCGLVD	148
Dd	283	LGNKLIRI-----YS-----DVERDYDGKFLEEINQVVSSISAVT-----TD	321
Oy	149	EKTVVVHTSQTPVPDPQGSTPLM	171
Dd	322	LQLAVMKTRMQPVGVKNFKFPRM	344

RESULT 10
G97438

glutamate-cysteine ligase precursor, chloroplast (gamma-glutamylcysteine synthetase) (Ga

C/Spectles: Agrobacterium tumefaciens
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C/Accession: G97438

R/Goodman, B.; Hinkle, G.; Gatungu, S.; Miller, N.; Blanchard, M.; Quirillo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2223-2328, 2001
S/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
S/Reference number: A97359; PMID:11743194

A/Accession: G97438
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-457 <KUR>
A/Cross-references: GB:AEO07869; PIDN:AAK86464.1; PID:G15155610; GSPDB:GN00169
C/Genetics:
A/Gene: AGR_C_1167
A/Map position: circular chromosome

```

Query Match 11                               8.6%; Score 78.5; DB 2; Length 457;
Glutamate-cysteine ligase gshl [imported] - Agrobacterium tumefaciens (strain C58, Dupont
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C;Accession: AB2657
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.;
Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; PMID:11743193
A;Accession: AB2657
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-457 <KDR>
A;Cross-references: GB:AF00688; PIDN:AAL1672.1; PID:G17739016; GSPDB:GN00186
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: gshl
A;Map position: circular chromosome

Query Match 33                               8.6%; Score 78.5; DB 2; Length 457;
Best local similarity 32.7%; Pred. No. 16;
Matches 33; Conservative 13; Mismatches 26; Indels 29; Gaps 6;

2  GGAAGPEKRPCCDLRCPETTSQALPAFTTEI---GEASEGPGA-----DEVQFAEAPAN 51
Db  325 GADGCPWRRIC-----ALPAPWGLLYNDEALDAADALRADMSFDE--VIALRN 371

Oy  52 ALPASEAAAV--QPVIGISQRY---RNNSKEKDLGLTG 86
Db  372 AVPAKGLAAEINGKPLGLIARQVLDSRTGLKNRRRLNNGE 412

RESULT 11
AB2657
glutamate-cysteine ligase gshl [imported] - Agrobacterium tumefaciens (strain C58, Dupont
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C;Accession: AB2657
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.;
Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; PMID:11743193
A;Accession: AB2657
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-457 <KDR>
A;Cross-references: GB:AF00688; PIDN:AAL1672.1; PID:G17739016; GSPDB:GN00186
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: gshl
A;Map position: circular chromosome

```

Qy	2	GEAYPEERPEDEDRCEPTTSQALPAFTTEI	-----QESSEGGCA-----	DEVOPAPAN	51
Db	325	GADGSPMRRC	-----ALPAFWGLLYNQEALDADALTDMSFDE	-VIALRN	371
Qy	52	ALPARSEMAAV	--OPVIGISORV----	RMSKKEKDKITG	86
Db	372	AVPAKGLAAETAGKPLUGIAKQVLDISTRIGAKNKRRLNGEG			412

RESULT 12
A32555
major merozoite surface antigen precursor - Plasmodium chabaudi adami (fragment)
C/Species: Plasmodium chabaudi adami
C/Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 11-Jan-2000
C/Accession: A32555
R/Lew, A.M.; Langford, C.J.; Anders, R.F.; Kemp, D.J.; Saul, A.; Fardoulis, C.; Geysen,
Proc. Natl. Acad. Sci. U.S.A. 86, 3768-3772, 1989
A/Title: A protective monoclonal antibody recognizes a linear epitope in the precursor to
A/Reference number: A32555; MUID:89264504; PMID:2471131
A/Accession: A32555
A/Molecule type: mRNA
A/Residues: 1-478 <LEM>
A/References: GB:J04568; NID:q160419; PID:g552206
C/Superfamily: major merozoite surface antigen
C/Keywords: glycoprotein; surface antigen
C/139,259/Binding site: carbohydrate (Asn) (covalent) #status predicted

```

OY      3  EAGVPEKPKPCDLCDEPTTSQALP-----AFTTEIOEASGPGADEV-----44
Db      266  EATQPESHAPAE-ATTETTTTAAEPETPERKEGASINXKSTSGCTARAPAEASTVEVPASPAT 324
          |||         |||||
OY      45  -----QVFAPANALPARSEAAAVPYITIGISQVRNASKKKD 81
          :|||       :|||
Db      325  PAAPSASSPAQAQPAAPAVTTSQPVPSGGSTNVCEGTOKAE 365
          :|||       :|||

```

```

RESULT 13
T22759
hypothetical protein F55H12.3 - Caenorhabditis elegans
C|Species: Caenorhabditis elegans
C|Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Aug-2000
C|Accession: T22759
R|Dobson, R.
submitted to the EMBL Data Library, October 1996
A|Reference number: Z19610
A|Accession: T22759
A|Status: preliminary; translated from GB/EMBL/DBJ
A|Molecule type: DNA
A|Residues: 1-2824 <WIL>
A|Cross-references: EMBL:Z81091; PIDs: CAB03143.1; GSPDB: GN00019; CESP: F55H12.3
A|Experimental source: clone F55H12
C|Genetics:
A|Gene: CESP: F55H12.3
A|Map position: 1
A|Intons: 87/1; 98/1; 126/2; 201/3; 343/3; 406/1; 576/3; 656/1; 825/3; 869/1; 909/1; 967/1; 1755/2; 1800/1; 1850/3; 1896/1; 2003/3; 2035/3; 2082/3; 2119/1; 2144/1; 2200/2; 2271/1
C|Superfamily: LDL receptor ligand-binding repeat homology
C|243-279/domain: LDL receptor ligand-binding repeat homology <LDL>

```

```

Query Match      8.6%; Score 78.5; DB 2; Length 2824;
Best Local Similarity 22.5%; Pred. No. 1.3e+02;
Matches 38; Conservative 19; Mismatches 69; Indels 43; Gaps 6;

QY    11 PCEDL-----RCPEPTTSQLPAPFTTEQEAASEGCGABEVVFAPANLPARSEAAA 61
       |||||
Db     2500 PCSDSSENATSI PVTYCOSTCSDALPTAGCNQLQONGS----MTRNCLP-----E 2638
               *

QY    62 VQPVIGISQRVRNMSKEKDDGLTGIVGIITMMVIAIIGAGIIIGYSKKRKDLKE-OH 120

```



```

Db 2639 VTFPPGNSNAIK-----VLGWVFCVLLILIVLVCFRQRIIAMPKCKDITSDNHQ 2689
QY 121 DQKVCERMQRIITPLSAFTNPICELIVDEKTVVHTSQTSPVDPOEGSTP 169
Db 2690 -----VALSHMDNATNRNEENQNPSTINYPPIPCAPIP 2725

```

RESULT 14
AD0760
diol dehydratase medium chain [imported] - *Salmonella enterica* subsp. *enterica* serovar 1
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typh
A:Note: this species has also been called *Salmonella typhi*
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C:Accession: AD0760
R:Parkhill, J.; Dougan, G.; James, K.D.; Thompson, N.R.; Pickard, D.; Wain, J.; Churcher
ch, T.; Connelton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar
S.; Mowle, S.; O'Garra, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov
A:Reference number: AB0502; EMBL:1677608
A:Accession: AD0760
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1,224 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD02402.1; PID:G16503272; GSPDB:GN00176
C:Genetics:
C:Gene: pduD

Query Match	8.5%;	Score 78;	DB 2;	Length 224;
Best Local Similarity	27.0%;	Pred. No. 7.9;		
Matches 34;	Conservative 15;	Mismatches 37;	Indels 40;	Gaps 6

```

OY 18 PETTSQALPATTEIQEASEGPADEVQVAPANALPARSEAAAVQPTIGISORVPMNSK 77
Db 37 PQTAAPAGDGFTEVEGARQGTQDEVI-----AVGPAFLAQTNIVGL 82

```

Qy 78 EKDLGTLGYL-----GITMVI-----IIAI-----GAGIILGSYKCKDL 116
| : | : || | :|| : |
Db 83 PKHGI-LREVIAGIEEGIRARVIRCFKSSDVAFAVEGNRLSGISIGI--QSMDT 137

QY	117	KEQHDQ	122
Db	138	TVIHQQ	143

RESULT 15
D98331
opdp protein (AF05387) [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C:Accession: D98331
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quirallo, B.; Goldman,
A.; Liu, F.; Wolman, C.; Allinger, M.; Doughty, D.; Scott, C.; Leppas, C.; Makelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97355; PMID:11743194
A:Accession: D98331
A:Status: Preliminary
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1267 <KOR>
A:Cross-references: GB:AE007870, FIDN:AAK90174.1, FID:G15160177, GSPDB:GN00170
C:Genetics:
A:Gene: AGR_L_3200
A:Map position: linear chromosome

Query Match	8.58;	Score 77.5;	DB 2;	Length 267;
Best Local Similarity	21.18;	Pred. No. 11;		
Matches 30;	Conservative 24;	Mismatches 43;	Indels 45;	Gaps 4

```

48 APANALPARSEAA-----AVQPVTGTSQVRNMSKKEKDLGTGLVGLITMMV 95
| | | | | : | : | | | : | :
QY

```

Db	53	ALNGLEPRARQVSGDLSWRSRSGLFTRPMPG-----	RDGLTIFQDPTATLNP	100
Qy	96	II-----IATGAGITLIGYVRYGKDLKEQ-----	HDQVYCRERMPRL	134
		----- ----- ----- ----- -----	----- ----- ----- ----- -----	
Db	101	VLTIEQVAGESGVRPHGLSGSWKGRGLRDRLLEVRVLPHPHLSATVPHQSGGQKQVAVI		160
Qy	135	PLSAFTNPCEIYDEKTVVHT		156
		----- ----- ----- ----- -----	----- ----- ----- ----- -----	
Db	161	AAATARRPALLIDENKTSALDT		182

```
Search completed: April 7, 2003, 09:19:23
Job time : 17.5135 secs
```


GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:24 ; Search time 8.18711 Seconds
(without alignments)
906.823 Million cell updates/sec

Title: US-10-057-951-2_COPY_85_263

Sequence: 1 SGBAGVPEKRPCEDLRCPET.....PVDQEGSTPLMQAGTPGA 179

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	86.5	9.4	820 1	SYL_CHLUPN
2	81	8.8	499 1	GSHR_PLAT7
3	81	8.8	499 1	GSHR_PLATK
4	80	8.7	1238 1	UAG2_HUMAN
5	77.5	8.5	1709 1	SN_HUMAN
6	76.5	8.4	425 1	TR16_RAT
7	76.5	8.4	501 1	GYG2_HUMAN
8	76	8.3	505 1	SPRO_SYNY3
9	75	8.2	595 1	TNR8_HUMAN
10	74.5	8.1	747 1	YRGP_ECOLI
11	73	8.0	260 1	MTKB_METEX
12	73	8.0	536 1	YENI_SCHPO
13	73	8.0	676 1	ICP0_HSVBU
14	73	8.0	698 1	EPG_DBIRA
15	72.5	7.9	400 1	BIAR_MOUSE
16	72.5	7.9	458 1	DESM_XENLA
17	72	7.9	646 1	MUS8_HUMAN
18	72	7.9	952 1	IF41_YEAST
19	72	7.9	1781 1	AKAC_HUMAN
20	71.5	7.8	145 1	ID54_METLA
21	71.5	7.8	755 1	RRE1_HUMAN
22	71.5	7.8	817 1	HUMB_MUSDO
23	70.5	7.7	241 1	MIAT_SARMT
24	70.5	7.7	272 1	IBP1_RAT
25	70.5	7.7	442 1	BNB_DROME
26	70.5	7.7	534 1	FM2_ACTNA
27	70.5	7.7	1001 1	IR2_SYNY3
28	70.5	7.7	1182 1	HAIR_MOUSE
29	70.5	7.7	2278 1	FAB1_YEAST
30	70	7.6	310 1	SUCA_DICDI
31	70	7.6	393 1	SDY1_RAT
32	70	7.6	652 1	CD93_HUMAN
33	70	7.6	1505 1	AT7B_SHEEP

34	69.5	7.6	241 1	MIAM_SARMT	008668 sarcocystis
35	69.5	7.6	272 1	IBP1_MOUSE	P47876 mus musculus
36	69.5	7.6	358 1	VCOM_ADP40	P48753 human adeno
37	69.5	7.6	388 1	NHAA_ECOLI	P13738 escherichia
38	69.5	7.6	513 1	ATPA_MYCFU	Q98q57 mycoplasma
39	69.5	7.6	579 1	PR17_HUMAN	O60508 homo sapien
40	69.5	7.6	626 1	PM17_MOUSE	O60536 mus musculu
41	69.5	7.6	1025 1	DPYD_PIG	Q28943 sus scrofa
42	69.5	7.5	322 1	SUCA_CAERL	P53596 caenorhabdi
43	69	7.5	382 1	NUSA_BACDH	Q9K474 bacillus ha
44	69	7.5	683 1	ERG2_TREPA	O83464 treponema p
45	69	7.5	766 1	METE_YEAST	P05694 saccharomyc

ALIGNMENTS

RESULT 1
ID SYL_CHLUPN STANDARD, PRT, 820 AA.
AC Q92930; Q9JQ86; 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Leucyl-tRNA synthetase (EC 6.1.1.4) (leucine--tRNA ligase) (LeuRS).
GN LEURS OR CPN0153 OR CP0618.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CW1029;
RX MEDLINE=9206606; PubMed=10192388;
RA Kaiman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S., Trachomatis.",
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.",
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Vetterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwyn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mohn and Chlamydia
RT pneumoniae AR39.",
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=UJ138;
RX MEDLINE=2030349; PubMed=10871362;
RA Shitara M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kohara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae UJ138
RT from Japan and CW1029 from USA.",
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [4]
RP CATALYTIC ACTIVITY: ATP + L-leucine + tRNA(Leu) = AMP +
RP diphosphate + L-leucyl-tRNA(Leu).
RN [5]
RP SUBCELLULAR LOCATION: Cytoplasmic.
RN [6]
RP SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
RN [7]
RP This SWISS-PROT entry is copyrighted. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RN modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
RN or send an email to license@isb-sib.ch).
RN EMBL: AB001602; AAD18306.1; --
RN EMBL: AB002219; AAF38433.1; --

DR EMBL: AP002545; BAA98363.1; -.
 DR TIGR: CP0618; -.
 DR InterPro: IPR002302; leu-cRNAseynta.
 DR InterPro: IPR002300; crna-synt_1a.
 DR InterPro: IPR001412; crna-synt_1.
 DR Pfam: PF00133; crna-synt_1; 1.
 DR PRINTS: PR00985; TRNASYNTHLEU.
 DR TIGRfam: TIGR00396; leu5 bact; 1.
 DR PROSITE: PS00178; AA-TRNA_LIGASE_I; 1.
 DR Anticacyl-cRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KM Anticacyl-cRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 FT SITE 40 51 "HIGH" REGION.
 FT BINDING 601 605 "RMSKS" REGION.
 FT ACT SITE 604 604 ATP (BY SIMILARITY).
 SQ SEQUENCE 820 AA; 93965 MW; 520369FC098F1926 CRC64;

Query Match 9.4%; Score 86.5; DB 1; Length 820;
 Best Local Similarity 24.1%; Pred. No. 2.9;

Matches 38; Conservative 28; Mismatches 49; Indels 43; Gaps 9;

QY 20 TTSGALPFTTEIQE-----ASEGPADEVQVFPANALPARSEAA-VQPVIGIS 69
 Db 235 TQESLFAFTTLDLLGVSFLVIAPEHPDLDSIV-----SEGRDEVTAVQESLRKS 288
 QY 70 QRVAMNS-KKKDLGLTGY----VLGITMVV-----ITAIAGAILGYSYKRGDKLEQ 119
 Db 289 ERDISSVKTGTGVTGNVAKHPIITGNLLPWISDVVLGVTGVWGV-----PA 339
 QY 120 HDQKVCEREMQRTLPASFTNPTCEIVDEKTVVHTS 157
 Db 340 HDER--DREFAEV-----FSLPIHEVIDNGVCVCHSN 369

RESULT 2

GSRR_PLAF7 STANDARD; PRT; 499 AA.
 ID GSRR_PLAF7
 AC 015770;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glutathione reductase (EC 1.6.4.2) (GR) (GRASE).
 GN GR1.
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gilbert R. T.-W., Walter R. D., Mueller S.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MAINTAIN HIGH LEVELS OF REDUCED GLUTATHIONE IN THE CYTOSOL. (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: NADPH + oxidized glutathione = NADP(+) + 2 glutathione.
 CC -1- COFACTOR: FAD (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- MISCELLANEOUS: THE ACTIVE SITE IS A REDOX-ACTIVE DISULFIDE BOND. OXIDOREDUCTASES CLASS-I.
 CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF027825; AAB84117.1; -.
 DR HSPG: P00390; 1ALG.
 DR InterPro: IPR001327; FAD_PYT_redox.
 DR InterPro: IPR000815; Hg_reductase.
 DR InterPro: IPR001100; Pyt_redox.
 DR InterPro: IPR000103; Pyridine_redox_2.

DR InterPro: IPR004099; pyt_redox_dim.
 DR Pfam: PF00070; pyt_redox_1.
 DR Pfam: PF02852; pyt_redox_dim; 1.
 DR PRINTS: PR00368; FADPDR.
 DR PRINTS: PR00945; HGRDPTASE.
 DR PRINTS: PR00411; PNDRTASEII.
 DR PRINTS: PR00469; PNDRTASEII.
 DR PRODOM: PD000139; FAD_PYT_redox; 2.
 DR PROSITE: PS00076; PYRIDINE_REDOX_1; 1.
 DR Redox-active center; Oxidoreductase; Flavoprotein; FAD; NADP.
 KM Redox-active center; Oxidoreductase; Flavoprotein; FAD; NADP.
 FT INT_MET 0 0
 FT NP_BIND 3 33 FAD (ADP PART) (POTENTIAL).
 FT DISULFID 39 44 REDOX-ACTIVE (BY SIMILARITY).
 FT NP_BIND 301 311 FAD (FLAVIN PART) (BY SIMILARITY).
 FT ACT SITE 484 484 BY SIMILARITY.
 SQ SEQUENCE 499 AA; 56288 MW; 43CCE0251E7B8244 CRC64;

Query Match 8.8%; Score 81; DB 1; Length 499;
 Best Local Similarity 24.5%; Pred. No. 5.3;

Matches 34; Conservative 25; Mismatches 64; Indels 16; Gaps 7;

QY 29 TTEIQASEGPADEVQVFPANALPARSEAAVQPVIGISQVR---NMSKEKDLGT 84
 Db 123 TKNNKNDNGPLNEF--ILEGRNIIAVGNKRVPPVAGIENTISSDFEYIKESKIGI 180
 QY 85 LGY-VLGITMMVITAIAGAILGYSYKRGDKLKHQ---KVCEREMQRTLPASFT 140
 Db 181 VSGSYIAVELINVIKRLG--IDSYIFARGNRIIRKFDSEVINVLNDKNNINIVTFA 237
 QY 141 NPTCEI--VDEKTVVHTS 157
 Db 238 D-VVEIKKVSQDKLSIHL 255

RESULT 3

GSRR_PLAFK STANDARD; PRT; 499 AA.
 ID GSRR_PLAFK
 AC 094655;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glutathione reductase (EC 1.6.4.2) (GR) (GRASE).
 GN GR2.
 OS Plasmodium falciparum (isolate K1 / Thailand).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5839;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=96370813; PubMed=87744709;
 RA Faerber P.M., Becker K., Mueller S.;
 RL "Molecular cloning and characterization of a putative glutathione reductase gene, the pGR2 gene, from Plasmodium falciparum."; Eur. J. Biochem. 239:655-661(1996).
 RN [2]
 RP SEQUENCE OF 1-15, AND CHARACTERIZATION.
 RX MEDLINE=96202957; PubMed=8611352;
 RA Kruttsch-Siegel R.L., Muller J.G., Lotzspeich F., Schirmer R.H.;
 RL "Glutathione reductase and glutamate dehydrogenase of Plasmodium falciparum, the causative agent of tropical malaria."; Eur. J. Biochem. 235:345-350(1996).
 CC -1- FUNCTION: MAINTAIN HIGH LEVELS OF REDUCED GLUTATHIONE IN THE CYTOSOL. (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: NADPH + oxidized glutathione = NADP(+) + 2 glutathione.
 CC -1- COFACTOR: FAD.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- MISCELLANEOUS: THE ACTIVE SITE IS A REDOX-ACTIVE DISULFIDE BOND. OXIDOREDUCTASES CLASS-I.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.embnet.org/licenses/embnet.html>)
 CC or send an email to license@ebi.ac.uk.
 CC -----
 DR EMBL: X93462: CAAG347.1; -.
 DR HSSP: P00390.1ALG.
 DR InterPro: IPR001327; FAD_pyr_redox.
 DR InterPro: IPR000815; Hg_reductase.
 DR InterPro: IPR001100; Pyr_redox.
 DR InterPro: IPR000103; Pyridine_redox_2.
 DR InterPro: IPR004099; pyr_redox_dim.
 DR Pfam: PF000070; pyr_redox; 1.
 DR Pfam: PF02852; pyr_redox; 1.
 DR PRINTS: PR00368; FADPNR.
 DR PRINTS: PR00945; HGRDPR.
 DR PRINTS: PR00411; PNDPRASRT.
 DR PRINTS: PR00469; PNDPRASRT.
 DR PRODOM: PD000139; FAD_pyr_redox; 2.
 DR PROSITE: PS00076; PYRIDINE_REDOX_1; 1.
 DR Redox-active center; Oxidoreductase; Flavoprotein; FAD; NADP.
 FT INT MET 0
 FT NP BIND 3 33 FAD (ADP PART) (POTENTIAL).
 FT DISULFID 39 44 REDOX-ACTIVE (BY SIMILARITY).
 FT NP BIND 301 311 FAD (FLAVIN PART) (BY SIMILARITY).
 FT ACT_SITE 484 484 BY SIMILARITY.
 SQ SEQUENCE 499 AA; 56430 MW; 6B239901ECA0B5F7 CRC64;
 Query Match 8.8%; Score 81; DB 1; Length 499;
 Best Local Similarity 24.5%; Pred. No. 5.3;
 Matches 34; Conservative 25; Mismatches 64; Indels 16; Gaps 7;
 QY 29 TTEIQEASGEGADVEVPAPANALPARSEAAVQPIGISOYR---MNSKEKDLGT 84
 DB 123 TKDNNKNDGPNLNE--ILEGRNLIAGVGMKPPVKGIENTISSDEFNIESKIGI 180
 QY 85 LGY-VLGITMNVIIAIGAGIIGLYSKRGKDLKEQHDQ---KYCEREMQRTLPISAF 140
 DB 181 VGSYIAVELINIVIKRG---ISYIPARGNRILIRKEDSVINVLNDMKNNINIVTFA 237
 QY 141 NPTCEI--VDEKTVVWTS 157
 DB 238 D-VWEIKKVDKNIHLHS 255
 RESULT 4
 JAG2 HUMAN
 ID JAG2 HUMAN STANDARD; PRT: 1238 AA.
 AC 09Y219; O9UE99; Q9Y6Q0; Q9Y6P9; Q9UNK8; Q9UE17;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Jagged 2 precursor (Jagged2) (HJ2).
 GN JAG2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (LONG ISOFORM).
 RP MEDLINE=97459705; PubMed=9315665;
 RA Luo B., Aster J.C., Hasseljian R.P., Kuo F., Sklar J.;
 RT "Isolation and functional analysis of a cDNA for human Jagged2, a gene
 RT encoding a ligand for the Notch1 receptor.";
 RL Mol. Cell. Biol. 17:6057-6067(1997).
 RL [2]
 RP SEQUENCE FROM N.A. (LONG ISOFORM).
 RP TISSUE=Fetal brain;
 RX MEDLINE=99180765; PubMed=10079256;
 RX Gray G.B., Mann R.S., Maltstadt E., Henriques D., Carcangiu M.-L.,
 RA Banks A., Leiman J., Ward D., Ish-Horowitz D., Artavanis-Tsakonas S.;
 RT "Human ligands of the Notch receptor.";

RL Am. J. Pathol. 154:785-794(1999).
 RN [3]
 RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
 RP TISSUE=Bone marrow;
 RX MEDLINE=20130121; PubMed=10662552;
 RA Deng Y., Madan A., Banta A.B., Friedman C., Trask B.J., Hood L.,
 RA Li L.;
 RT "Characterization, chromosomal localization, and the complete 30-kb
 RT DNA sequence of the human Jagged2 (JAG2) gene.";
 RL Genomics 63:133-138(2000).
 RN [4]
 RP SEQUENCE OF 17-1238 FROM N.A. (LONG ISOFORM).
 RP TISSUE=Heart;
 RX MEDLINE=98145947; PubMed=9486542;
 RA Valasechi C., Ghezzi C., Ballabio A., Rugarli E.L.;
 RT "JAGGED2: a putative Notch ligand expressed in the apical ectodermal
 RT ridge and in sites of epithelial-mesenchymal interactions.";
 RL Mech. Dev. 69:203-207(1997).
 CC -1- FUNCTION: PUTATIVE NOTCH LIGAND INVOLVED IN THE MEDIATION OF NOTCH
 CC SIGNALING. INVOLVED IN LIMB DEVELOPMENT (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM/HJAG2.DEL-86; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN HEART, PLACENTA AND SKELETAL
 CC MUSCLE AND TO A LESSER EXTENT IN PANCREAS. VERY LOW EXPRESSION IN
 CC BRAIN, LUNG, LIVER AND KIDNEY.
 CC -1- DISEASE: MAY BE ASSOCIATED TO USHER SYNDROME TYPE 1A (USH1A)
 CC WHICH DESCRIBES A CONGENITAL SENSOR DEAFNESS ASSOCIATED WITH
 CC RETINITIS PIGMENTOSA AND FEEBLE-MINDEDNESS.
 CC -1- SIMILARITY: CONTAINS 16 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 DSL DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.embnet.org/licenses/embnet.html>)
 CC or send an email to license@ebi.ac.uk.
 CC -----
 DR EMBL: AF020201; AAB71189.1; -.
 DR EMBL: AF003521; AAB61285.1; -.
 DR EMBL: AF029778; AAB84215.1; -.
 DR EMBL: AF029779; AAB84216.1; -.
 DR EMBL: AF111170; AAD15562.1; -.
 DR EMBL: Y14330; CAAT7406.1; -.
 DR HSSP: P00743; ICGF.
 DR GeneW: HGNC:6189; JAG2.
 DR MIM: 602570; -.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR001774; DSL.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000742; EGF_2.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF00008; EGF_14.
 DR Pfam: PF01414; DSL; 1.
 DR PRINTS: PR00010; EGFRLCOO.
 DR PRINTS: PR00011; EGFRLAMININ.
 DR SMART: SM00051; DSL; 1.
 DR SMART: SM00179; EGF_CA; 9.
 DR SMART: SM00001; EGF_1like; 6.
 DR SMART: SM00214; VWF_1.
 DR PROSITE: PS00010; ASX_HYDROXYL; 10.
 DR PROSITE: PS00022; EGF_1; 16.
 DR PROSITE: PS01186; EGF_2; 12.
 DR PROSITE: PS01187; EGF_CA; 7.
 DR Calcium-binding; EGF-like domain; Glycoprotein; Developmental protein;
 KW Repeat; Signal; Transmembrane; Alternative splicing.
 FT SIGNAL 1 26
 FT CHAIN 27 1238 JAGGED 2.
 FT DOMAIN 27 1080 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1081 1105 POTENTIAL.

FT DOMAIN 1106 1238 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 178 240 DSL.
FT DOMAIN 241 274 EGF-LIKE 1.
FT DOMAIN 275 305 EGF-LIKE 2.
FT DOMAIN 307 345 EGF-LIKE 3.
FT DOMAIN 347 383 EGF-LIKE 4.
FT DOMAIN 385 421 EGF-LIKE 5.
FT DOMAIN 423 459 EGF-LIKE 6.
FT DOMAIN 461 496 EGF-LIKE 7.
FT DOMAIN 498 534 EGF-LIKE 8.
FT DOMAIN 536 572 EGF-LIKE 9.
FT DOMAIN 574 634 EGF-LIKE 10 (ATYPICAL).
FT DOMAIN 636 672 EGF-LIKE 11.
FT DOMAIN 674 710 EGF-LIKE 12.
FT DOMAIN 712 748 EGF-LIKE 13.
FT DOMAIN 751 787 EGF-LIKE 14.
FT DOMAIN 789 825 EGF-LIKE 15.
FT DOMAIN 827 863 EGF-LIKE 16.
FT DOMAIN 863 899 EGF-LIKE 17.
FT DOMAIN 899 935 EGF-LIKE 18.
FT DOMAIN 937 973 EGF-LIKE 19.
FT DOMAIN 975 1011 EGF-LIKE 20.
FT DOMAIN 1013 1049 EGF-LIKE 21.
FT DOMAIN 1051 1087 EGF-LIKE 22.
FT DOMAIN 1089 1125 EGF-LIKE 23.
FT DOMAIN 1127 1163 EGF-LIKE 24.
FT DOMAIN 1165 1201 EGF-LIKE 25.
FT DOMAIN 1203 1239 EGF-LIKE 26.
FT DOMAIN 1241 1277 EGF-LIKE 27.
FT DOMAIN 1279 1315 EGF-LIKE 28.
FT DOMAIN 1317 1353 EGF-LIKE 29.
FT DOMAIN 1355 1391 EGF-LIKE 30.
FT DOMAIN 1393 1429 EGF-LIKE 31.
FT DOMAIN 1431 1467 EGF-LIKE 32.
FT DOMAIN 1469 1505 EGF-LIKE 33.
FT DOMAIN 1507 1543 EGF-LIKE 34.
FT DOMAIN 1545 1581 EGF-LIKE 35.
FT DOMAIN 1583 1619 EGF-LIKE 36.
FT DOMAIN 1621 1657 EGF-LIKE 37.
FT DOMAIN 1659 1695 EGF-LIKE 38.
FT DOMAIN 1697 1733 EGF-LIKE 39.
FT DOMAIN 1735 1771 EGF-LIKE 40.
FT DOMAIN 1773 1809 EGF-LIKE 41.
FT DOMAIN 1811 1847 EGF-LIKE 42.
FT DOMAIN 1849 1885 EGF-LIKE 43.
FT DOMAIN 1887 1923 EGF-LIKE 44.
FT DOMAIN 1925 1961 EGF-LIKE 45.
FT DOMAIN 1963 1999 EGF-LIKE 46.
FT DOMAIN 2001 2037 EGF-LIKE 47.
FT DOMAIN 2039 2075 EGF-LIKE 48.
FT DOMAIN 2077 2113 EGF-LIKE 49.
FT DOMAIN 2115 2151 EGF-LIKE 50.
FT DOMAIN 2153 2189 EGF-LIKE 51.
FT DOMAIN 2191 2227 EGF-LIKE 52.
FT DOMAIN 2229 2265 EGF-LIKE 53.
FT DOMAIN 2267 2303 EGF-LIKE 54.
FT DOMAIN 2305 2341 EGF-LIKE 55.
FT DOMAIN 2343 2379 EGF-LIKE 56.
FT DOMAIN 2381 2417 EGF-LIKE 57.
FT DOMAIN 2419 2455 EGF-LIKE 58.
FT DOMAIN 2457 2493 EGF-LIKE 59.
FT DOMAIN 2495 2531 EGF-LIKE 60.
FT DOMAIN 2533 2569 EGF-LIKE 61.
FT DOMAIN 2571 2607 EGF-LIKE 62.
FT DOMAIN 2609 2645 EGF-LIKE 63.
FT DOMAIN 2647 2683 EGF-LIKE 64.
FT DOMAIN 2685 2721 EGF-LIKE 65.
FT DOMAIN 2723 2759 EGF-LIKE 66.
FT DOMAIN 2761 2797 EGF-LIKE 67.
FT DOMAIN 2799 2835 EGF-LIKE 68.
FT DOMAIN 2837 2873 EGF-LIKE 69.
FT DOMAIN 2875 2911 EGF-LIKE 70.
FT DOMAIN 2913 2949 EGF-LIKE 71.
FT DOMAIN 2951 2987 EGF-LIKE 72.
FT DOMAIN 2989 3025 EGF-LIKE 73.
FT DOMAIN 3027 3063 EGF-LIKE 74.
FT DOMAIN 3065 3101 EGF-LIKE 75.
FT DOMAIN 3103 3139 EGF-LIKE 76.
FT DOMAIN 3141 3177 EGF-LIKE 77.
FT DOMAIN 3179 3215 EGF-LIKE 78.
FT DOMAIN 3217 3253 EGF-LIKE 79.
FT DOMAIN 3255 3291 EGF-LIKE 80.
FT DOMAIN 3293 3329 EGF-LIKE 81.
FT DOMAIN 3331 3367 EGF-LIKE 82.
FT DOMAIN 3369 3405 EGF-LIKE 83.
FT DOMAIN 3407 3443 EGF-LIKE 84.
FT DOMAIN 3445 3481 EGF-LIKE 85.
FT DOMAIN 3483 3519 EGF-LIKE 86.
FT DOMAIN 3521 3557 EGF-LIKE 87.
FT DOMAIN 3559 3595 EGF-LIKE 88.
FT DOMAIN 3597 3633 EGF-LIKE 89.
FT DOMAIN 3635 3671 EGF-LIKE 90.
FT DOMAIN 3673 3709 EGF-LIKE 91.
FT DOMAIN 3711 3747 EGF-LIKE 92.
FT DOMAIN 3749 3785 EGF-LIKE 93.
FT DOMAIN 3787 3823 EGF-LIKE 94.
FT DOMAIN 3825 3861 EGF-LIKE 95.
FT DOMAIN 3863 3899 EGF-LIKE 96.
FT DOMAIN 3901 3937 EGF-LIKE 97.
FT DOMAIN 3939 3975 EGF-LIKE 98.
FT DOMAIN 3977 4013 EGF-LIKE 99.
FT DOMAIN 4015 4051 EGF-LIKE 100.
FT DOMAIN 4053 4089 EGF-LIKE 101.
FT DOMAIN 4091 4127 EGF-LIKE 102.
FT DOMAIN 4129 4165 EGF-LIKE 103.
FT DOMAIN 4167 4203 EGF-LIKE 104.
FT DOMAIN 4205 4241 EGF-LIKE 105.
FT DOMAIN 4243 4279 EGF-LIKE 106.
FT DOMAIN 4281 4317 EGF-LIKE 107.
FT DOMAIN 4319 4355 EGF-LIKE 108.
FT DOMAIN 4357 4393 EGF-LIKE 109.
FT DOMAIN 4395 4431 EGF-LIKE 110.
FT DOMAIN 4433 4469 EGF-LIKE 111.
FT DOMAIN 4471 4507 EGF-LIKE 112.
FT DOMAIN 4509 4545 EGF-LIKE 113.
FT DOMAIN 4547 4583 EGF-LIKE 114.
FT DOMAIN 4585 4621 EGF-LIKE 115.
FT DOMAIN 4623 4659 EGF-LIKE 116.
FT DOMAIN 4661 4697 EGF-LIKE 117.
FT DOMAIN 4699 4735 EGF-LIKE 118.
FT DOMAIN 4737 4773 EGF-LIKE 119.
FT DOMAIN 4775 4811 EGF-LIKE 120.
FT DOMAIN 4813 4849 EGF-LIKE 121.
FT DOMAIN 4851 4887 EGF-LIKE 122.
FT DOMAIN 4889 4925 EGF-LIKE 123.
FT DOMAIN 4927 4963 EGF-LIKE 124.
FT DOMAIN 4965 5001 EGF-LIKE 125.
FT DOMAIN 5003 5039 EGF-LIKE 126.
FT DOMAIN 5041 5077 EGF-LIKE 127.
FT DOMAIN 5079 5115 EGF-LIKE 128.
FT DOMAIN 5117 5153 EGF-LIKE 129.
FT DOMAIN 5155 5191 EGF-LIKE 130.
FT DOMAIN 5193 5229 EGF-LIKE 131.
FT DOMAIN 5231 5267 EGF-LIKE 132.
FT DOMAIN 5269 5305 EGF-LIKE 133.
FT DOMAIN 5307 5343 EGF-LIKE 134.
FT DOMAIN 5345 5381 EGF-LIKE 135.
FT DOMAIN 5383 5419 EGF-LIKE 136.
FT DOMAIN 5421 5457 EGF-LIKE 137.
FT DOMAIN 5459 5495 EGF-LIKE 138.
FT DOMAIN 5497 5533 EGF-LIKE 139.
FT DOMAIN 5535 5571 EGF-LIKE 140.
FT DOMAIN 5573 5609 EGF-LIKE 141.
FT DOMAIN 5611 5647 EGF-LIKE 142.
FT DOMAIN 5649 5685 EGF-LIKE 143.
FT DOMAIN 5687 5723 EGF-LIKE 144.
FT DOMAIN 5725 5761 EGF-LIKE 145.
FT DOMAIN 5763 5799 EGF-LIKE 146.
FT DOMAIN 5801 5837 EGF-LIKE 147.
FT DOMAIN 5839 5875 EGF-LIKE 148.
FT DOMAIN 5877 5913 EGF-LIKE 149.
FT DOMAIN 5915 5951 EGF-LIKE 150.
FT DOMAIN 5953 5989 EGF-LIKE 151.
FT DOMAIN 5991 6027 EGF-LIKE 152.
FT DOMAIN 6029 6065 EGF-LIKE 153.
FT DOMAIN 6067 6103 EGF-LIKE 154.
FT DOMAIN 6105 6141 EGF-LIKE 155.
FT DOMAIN 6143 6179 EGF-LIKE 156.
FT DOMAIN 6181 6217 EGF-LIKE 157.
FT DOMAIN 6219 6255 EGF-LIKE 158.
FT DOMAIN 6257 6293 EGF-LIKE 159.
FT DOMAIN 6295 6331 EGF-LIKE 160.
FT DOMAIN 6333 6369 EGF-LIKE 161.
FT DOMAIN 6371 6407 EGF-LIKE 162.
FT DOMAIN 6409 6445 EGF-LIKE 163.
FT DOMAIN 6447 6483 EGF-LIKE 164.
FT DOMAIN 6485 6521 EGF-LIKE 165.
FT DOMAIN 6523 6559 EGF-LIKE 166.
FT DOMAIN 6561 6597 EGF-LIKE 167.
FT DOMAIN 6599 6635 EGF-LIKE 168.
FT DOMAIN 6637 6673 EGF-LIKE 169.
FT DOMAIN 6675 6711 EGF-LIKE 170.
FT DOMAIN 6713 6749 EGF-LIKE 171.
FT DOMAIN 6751 6787 EGF-LIKE 172.
FT DOMAIN 6789 6825 EGF-LIKE 173.
FT DOMAIN 6827 6863 EGF-LIKE 174.
FT DOMAIN 6865 6901 EGF-LIKE 175.
FT DOMAIN 6903 6939 EGF-LIKE 176.
FT DOMAIN 6941 6977 EGF-LIKE 177.
FT DOMAIN 6979 7015 EGF-LIKE 178.
FT DOMAIN 7017 7053 EGF-LIKE 179.
FT DOMAIN 7055 7091 EGF-LIKE 180.
FT DOMAIN 7093 7129 EGF-LIKE 181.
FT DOMAIN 7131 7167 EGF-LIKE 182.
FT DOMAIN 7169 7205 EGF-LIKE 183.
FT DOMAIN 7207 7243 EGF-LIKE 184.
FT DOMAIN 7245 7281 EGF-LIKE 185.
FT DOMAIN 7283 7319 EGF-LIKE 186.
FT DOMAIN 7321 7357 EGF-LIKE 187.
FT DOMAIN 7359 7395 EGF-LIKE 188.
FT DOMAIN 7397 7433 EGF-LIKE 189.
FT DOMAIN 7435 7471 EGF-LIKE 190.
FT DOMAIN 7473 7509 EGF-LIKE 191.
FT DOMAIN 7511 7547 EGF-LIKE 192.
FT DOMAIN 7549 7585 EGF-LIKE 193.
FT DOMAIN 7587 7623 EGF-LIKE 194.
FT DOMAIN 7625 7661 EGF-LIKE 195.
FT DOMAIN 7663 7699 EGF-LIKE 196.
FT DOMAIN 7701 7737 EGF-LIKE 197.
FT DOMAIN 7739 7775 EGF-LIKE 198.
FT DOMAIN 7777 7813 EGF-LIKE 199.
FT DOMAIN 7815 7851 EGF-LIKE 200.
FT DOMAIN 7853 7889 EGF-LIKE 201.
FT DOMAIN 7891 7927 EGF-LIKE 202.
FT DOMAIN 7929 7965 EGF-LIKE 203.
FT DOMAIN 7967 8003 EGF-LIKE 204.
FT DOMAIN 8005 8041 EGF-LIKE 205.
FT DOMAIN 8043 8079 EGF-LIKE 206.
FT DOMAIN 8081 8117 EGF-LIKE 207.
FT DOMAIN 8119 8155 EGF-LIKE 208.
FT DOMAIN 8157 8193 EGF-LIKE 209.
FT DOMAIN 8195 8231 EGF-LIKE 210.
FT DOMAIN 8233 8269 EGF-LIKE 211.
FT DOMAIN 8271 8307 EGF-LIKE 212.
FT DOMAIN 8309 8345 EGF-LIKE 213.
FT DOMAIN 8347 8383 EGF-LIKE 214.
FT DOMAIN 8385 8421 EGF-LIKE 215.
FT DOMAIN 8423 8459 EGF-LIKE 216.
FT DOMAIN 8461 8497 EGF-LIKE 217.
FT DOMAIN 8499 8535 EGF-LIKE 218.
FT DOMAIN 8537 8573 EGF-LIKE 219.
FT DOMAIN 8575 8611 EGF-LIKE 220.
FT DOMAIN 8613 8649 EGF-LIKE 221.
FT DOMAIN 8651 8687 EGF-LIKE 222.
FT DOMAIN 8689 8725 EGF-LIKE 223.
FT DOMAIN 8727 8763 EGF-LIKE 224.
FT DOMAIN 8765 8801 EGF-LIKE 225.
FT DOMAIN 8803 8839 EGF-LIKE 226.
FT DOMAIN 8841 8877 EGF-LIKE 227.
FT DOMAIN 8879 8915 EGF-LIKE 228.
FT DOMAIN 8917 8953 EGF-LIKE 229.
FT DOMAIN 8955 8991 EGF-LIKE 230.
FT DOMAIN 8993 9029 EGF-LIKE 231.
FT DOMAIN 9031 9067 EGF-LIKE 232.
FT DOMAIN 9069 9105 EGF-LIKE 233.
FT DOMAIN 9107 9143 EGF-LIKE 234.
FT DOMAIN 9145 9181 EGF-LIKE 235.
FT DOMAIN 9183 9219 EGF-LIKE 236.
FT DOMAIN 9221 9257 EGF-LIKE 237.
FT DOMAIN 9259 9295 EGF-LIKE 238.
FT DOMAIN 9297 9333 EGF-LIKE 239.
FT DOMAIN 9335 9371 EGF-LIKE 240.
FT DOMAIN 9373 9409 EGF-LIKE 241.
FT DOMAIN 9411 9447 EGF-LIKE 242.
FT DOMAIN 9449 9485 EGF-LIKE 243.
FT DOMAIN 9487 9523 EGF-LIKE 244.
FT DOMAIN 9525 9561 EGF-LIKE 245.
FT DOMAIN 9563 9599 EGF-LIKE 246.
FT DOMAIN 9601 9637 EGF-LIKE 247.
FT DOMAIN 9639 9675 EGF-LIKE 248.
FT DOMAIN 9677 9713 EGF-LIKE 249.
FT DOMAIN 9715 9751 EGF-LIKE 250.
FT DOMAIN 9753 9789 EGF-LIKE 251.
FT DOMAIN 9791 9827 EGF-LIKE 252.
FT DOMAIN 9829 9865 EGF-LIKE 253.
FT DOMAIN 9867 9903 EGF-LIKE 254.
FT DOMAIN 9905 9941 EGF-LIKE 255.
FT DOMAIN 9943 9979 EGF-LIKE 256.
FT DOMAIN 9981 10000 EGF-LIKE 257.

FT CONFLICT 119 119 A -> P (IN REF. 1).
FT CONFLICT 129 129 L -> F (IN REF. 1).
FT CONFLICT 384 384 L -> SA (IN REF. 4).
FT CONFLICT 424 426 ANE -> VND (IN REF. 1).
FT CONFLICT 501 501 K -> E (IN REF. 3).
FT CONFLICT 1235 1235 A -> V (IN REF. 2).
SQ SEQUENCE 1238 AA; 133364 MW; DA487BEC619933CE CRC64;
Query Match 8.7%; Score 80; DB 1; Length 1238;
Best Local Similarity 22.4%; Pred. No. 18;
Matches 55; Conservative 20; Mismatches 77; Indels 94; Gaps 11;
OY 2 GEAGV-PEKPC-----EDLRC-----PETTS-----QALPAFTTEIQ 33
Dy 947 GEGCABEPSPPCPRSGHDNNCARLTLEHNDHVGQTTVAGCSGIRSLPATRAVAR 1006
OY 34 E-----ASEPGADEVOY-PAPANALPAS--EAAVQVIGISQVRNNSKEKD 81
Dy 1007 DRLVLLCDRASGASAVEAVSFPARDLPDSLLIGAAHAIVALTOR----- 1056
OY 82 LGTLGVIGITMVIIAIGGILG-----YSYRGKDLKE 118
Dy 1057 -GNSSLIAVEVEVETVVGSGSTGLVYLGCASFVIMLACVVLQVMTTRRRK----- 1111
OY 119 QHDKVCEREMQRT-----LPLSAFTNPTCEIYDEKTVVHTSQRPVDPQEGSTP 169
Dy 1112 -----ERERSRLPRESANNQWAPLPIRNPIERPGKHVDVLYCKNFTPPRRADEA 1164
OY 170 LMGQAG 175
Dy 1165 LFGPAG 1170
RESULT 5
SN_HUMAN STANDARD; PRT; 1709 AA.
ID SN_HUMAN
AC Q9B222; Q9H1H6; Q9H1H7; Q9H1L7; Q9G2S5;
DT 16-OCT-2001 (Ref. 40, Created)
DT 16-OCT-2001 (Ref. 40, Last sequence update)
DT 15-JUN-2002 (Ref. 41, Last annotation update)
DE Sialoadhesin precursor (Sialic acid binding Ig-like lectin-1) (Siglec-1) (CD169 antigen).
GN SN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Hartnell A., Steel J., Turley H., Jones M., Jackson D., Crocker P.R.;
RT "Characterisation of human sialoadhesin (siglec-1), a sialic acid
RT binding receptor expressed on resident and inflammatory macrophage
RT populations",
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavriles G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deaman R., Dhani P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Frazer A., French L., Garner P.,
RA Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Leharstaino M.H., Leverha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McCormachie L.J., Mcclay K., McMurtry A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,

RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.B., Sehra H.K., Showkhen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Symamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin R., Wall M., Wallis J.W.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.,
 RL "The DNA sequence and comparative analysis of human chromosome 20.",
 RL Nature 414:865-871(2001).
 RP [3]
 RP SEQUENCE OF 733-1709 FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE-Spleen;
 RA Ohara O., Nagase T., Kikuno R., Okumura K.;
 RT "The nucleotide sequence of a long cDNA clone isolated from human
 spleen.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MACROPHAGE-RESTRICTED ADHESION MOLECULE THAT MEDIATES
 CC -1- STATIC-ACID DEPENDENT BINDING TO CELLS.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (ISOFORM 1) AND
 CC -1- SOLUBLE (ISOFORM 2).
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
 CC -1- produced by alternative splicing.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 16 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- DATABASE: NAME=PROW; NOTE=PROW 2:18-22(2001).
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/guide/985165905.g.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF230073; AAK0757.1; -;
 DR EMBL: AL109804; CAC17543.1; -;
 DR EMBL: AL109804; CAC17542.1; -;
 DR EMBL: AK024462; BAB15752.1; -;
 DR EMBL: AK024459; BAB15749.1; -;
 DR EMBL: AK024479; BAB15769.1; -;
 DR HSSP: O62230; 10FO.
 DR Genew: HGNC:11127; SN.
 DR MIM: 600751; -;
 DR InterPro: IPR003599; Ig.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003598; Ig_C2.
 DR InterPro: IPR003600; Ig_Like.
 DR Pfam: PF00047; Ig; 14.
 DR SMART: SM00409; Ig; 17.
 DR SMART: SM00410; Ig_Like; 1.
 DR SMART: SM00408; IGc2; 14.
 KW Signal; Glycoprotein; Cell adhesion; Immunoglobulin domain; Repeat;
 KW Transmembrane; Alternative splicing; Antigen.
 FT SIGNAL 1 19
 FT CHAIN 20 1709
 FT DOMAIN 20 1641
 FT TRANSMEM 1642 1662
 FT DOMAIN 1663 1709
 FT DOMAIN 20 136
 FT DOMAIN 153 224
 FT DOMAIN 255 312
 FT DOMAIN 339 397
 FT DOMAIN 426 498
 FT DOMAIN 524 582
 FT DOMAIN 617 696
 FT DOMAIN 722 781
 FT DOMAIN 810 883
 FT DOMAIN 909 967
 FT DOMAIN 998 1074
 FT IG-LIKE C2-TYPE DOMAIN 1.
 FT IG-LIKE C2-TYPE DOMAIN 2.
 FT IG-LIKE C2-TYPE DOMAIN 3.
 FT IG-LIKE C2-TYPE DOMAIN 4.
 FT IG-LIKE C2-TYPE DOMAIN 5.
 FT IG-LIKE C2-TYPE DOMAIN 6.
 FT IG-LIKE C2-TYPE DOMAIN 7.
 FT IG-LIKE C2-TYPE DOMAIN 8.
 FT IG-LIKE C2-TYPE DOMAIN 9.
 FT IG-LIKE C2-TYPE DOMAIN 10.

FT DOMAIN 1100 1156 IG-LIKE C2-TYPE DOMAIN 11.
 FT DOMAIN 1186 1248 IG-LIKE C2-TYPE DOMAIN 12.
 FT DOMAIN 1274 1331 IG-LIKE C2-TYPE DOMAIN 13.
 FT DOMAIN 1360 1432 IG-LIKE C2-TYPE DOMAIN 14.
 FT DOMAIN 1458 1518 IG-LIKE C2-TYPE DOMAIN 15.
 FT DOMAIN 1547 1620 IG-LIKE C2-TYPE DOMAIN 16.
 FT DISULFID 41 98
 FT DISULFID 160 217
 FT DISULFID 262 305
 FT DISULFID 346 390
 FT DISULFID 433 491
 FT DISULFID 531 575
 FT DISULFID 624 689
 FT DISULFID 729 774
 FT DISULFID 817 876
 FT DISULFID 916 960
 FT DISULFID 1005 1067
 FT DISULFID 1107 1149
 FT DISULFID 1193 1241
 FT DISULFID 1281 1324
 FT DISULFID 1367 1425
 FT DISULFID 1465 1511
 FT DISULFID 1554 1613
 FT CARBOHYD 159 159
 FT CARBOHYD 265 265
 FT CARBOHYD 339 339
 FT CARBOHYD 499 499
 FT CARBOHYD 697 697
 FT CARBOHYD 726 726
 FT CARBOHYD 730 730
 FT CARBOHYD 741 741
 FT CARBOHYD 886 886
 FT CARBOHYD 1104 1104
 FT CARBOHYD 1138 1138
 FT CARBOHYD 1251 1251
 FT CARBOHYD 1462 1462
 FT CARBOHYD 1476 1476
 FT VARSPLIC 1632 1709
 FT CONFLICT 1349 1349
 FT CONFLICT 1519 1519
 SQ SEQUENCE 1709 AA; 182624 MW; 587C7CCA0B789A6D CR664;
 Query Match 8.5%; Score 77.5; DB 1; Length 1709;
 Best Local Similarity 19.1%; Pred. No. 44;
 Matches 31; Conservative 32; Mismatches 56; Indels 43; Gaps 6;
 QY 22 SQALPAFTTET-----QEASRGGADEVYFPANALPARSEA----- 59
 DB 1558 SEPLASTLTLHGSRLVASSQPGGAFAEPHIVLASPNLRVDIELRPSDGEYICASN 1617
 QY 60 ----AAVGPVIGISQVRNRNSKEKKDLGTLGVLSITMMVITIAIGAGIILGYSKRGKD 115
 DB 1618 VLGASASTTYFGVRLRHLHQOQ-----LLMVLGLVGLLLLLGLGAC--YTWRR--- 1667
 QY 116 LKEQHDKVQKCERMQRITLPLSAFTNPCEIYDEKTVVHTS 157
 DB 1668 -----RVCYKQSMGNSVEM-AFOKETTQLIDPDPAACETS 1702
 RESULT 6
 TR16 RAT
 ID TR16 RAT STANDARD; PRT; 425 AA.
 AC P07174;
 DT 01-APR-1988 (rel. 07, Created)
 DT 01-APR-1988 (rel. 07, Last sequence update)
 DT 15-JUN-2002 (rel. 41, Last annotation update)
 DE Tumor necrosis factor receptor superfamily member 16 precursor (Low-
 DE affinity nerve growth factor receptor) (NGF receptor) (p80-LINGFR)
 DE (p75 ICD) (Low affinity neurotrophin receptor p75NTR).
 OS NGFR OR TNFRSF16.
 OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87115859; PubMed=3027580;
 RA Radeke M.J., Misko T.P., Hsu C., Herzenberg L.A., Shooter E.M.;
 RT "Gene transfer and molecular cloning of the rat nerve growth factor
 receptor.";
 RL Nature 325:593-597(1987).
 RN [2]
 RP SEQUENCE OF 1-22 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=93077038; PubMed=1446821;
 RA Meteis M., Timmusk T., Allikmets R., Saarma M., Persson H.;
 RT "Regulatory elements and transcriptional regulation by testosterone
 and retinoic acid of the rat nerve growth factor receptor promoter.";
 RL Gene 121:247-254(1992).
 RN [3]
 RP STRUCTURE BY NMR OF 334-418.
 RX MEDLINE=97449145; PubMed=9305641;
 RA Liepinsh E., Ilag L.L., Otting G., Ibanez C.F.;
 RT "NMR structure of the death domain of the p75 neurotrophin receptor.";
 RL EMBO J. 16:4999-5005(1997).
 CC -1- FUNCTION: Low affinity receptor which can bind to NGF, BDNF, NT-3,
 and NT-4. Can mediate cell survival as well as cell death of
 neural cells.
 CC -1- SUBUNIT: Homodimer; disulfide-linked. Interacts with p75NTR-
 associated cell death executor.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- PTM: N- AND O-glycosylated.
 CC -1- PTM: Phosphorylated on serine residues.
 CC -1- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.
 CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as there is no content in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X05137; CAA28783.1; -;
 DR EMBL; X61269; -; NOT_ANNOTATED_CDS.
 DR PIR; A26431; A26431.
 DR PDB; INGR; 29-JUL-97.
 DR InterPro; IPR000488; Death.
 DR InterPro; IPR001368; TNFR_c6.
 DR Pfam; PF00020; TNFR_c6; 4.
 DR Pfam; PF00531; death; 1.
 DR SMART; SM00208; TNFR; 3.
 DR SMART; SM00208; TNFR; 3.
 DR PROSITE; PS00652; TNFR_NGFR_1; 3.
 DR PROSITE; PS00652; TNFR_NGFR_2; 4.
 DR PROSITE; PS50017; DEATH_DOMAIN; 1.
 KW Repeat; Apoptosis; Neurogenesis; Transmembrane; Glycoprotein;
 KW Repeat; Phosphorylation; Signal; 3D-structure.
 FT SIGNAL 1 29
 FT CHAIN 30 425
 FT TUMOR NECROSIS FACTOR RECEPTOR
 FT SUPERFAMILY MEMBER 16.
 FT TRANSMEM 252 251
 FT TRANSMEM 273
 FT DOMAIN 274 425
 FT CYTOPLASMIC (POTENTIAL).
 FT REPEAT 32 65
 FT REPEAT 67 108
 FT REPEAT 109 147
 FT REPEAT 149 189
 FT REPEAT 198 249
 FT DOMAIN 354 419
 FT DEATH.
 FT SER/THR-RICH.
 FT DISULFID 33 44
 FT DISULFID 45 58
 FT DISULFID 48 65
 FT BY SIMILARITY.
 FT BY SIMILARITY.
 FT BY SIMILARITY.

FT DISULFID 68 84
 FT BY SIMILARITY.
 FT DISULFID 87 100
 FT BY SIMILARITY.
 FT DISULFID 90 108
 FT BY SIMILARITY.
 FT DISULFID 110 123
 FT BY SIMILARITY.
 FT DISULFID 126 139
 FT BY SIMILARITY.
 FT DISULFID 129 147
 FT BY SIMILARITY.
 FT DISULFID 150 165
 FT BY SIMILARITY.
 FT DISULFID 168 181
 FT BY SIMILARITY.
 FT DISULFID 171 189
 FT BY SIMILARITY.
 FT CARBOHYD 61 61
 FT CARBOHYD 71 71
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 425 AA; 45432 MW; B2E152D94D827F8 CRC64;
 Query Match 8.4%; Score 76.5; DB 1; Length 425;
 Best Local Similarity 24.4%; Pred. No. 12;
 Matches 42; Conservative 23; Mismatches 90; Indels 17; Gaps 5;
 QY 3 EAGVPEKPCEDLRCPETSOALPAFTTEIQEASSEGGADEVQVAFANLPAARSEAAV 62
 Db 175 EQRLRECPMADACEIEIGRWIPRSTPPGSDSTABSTOEPEV-PPEQDLVPSVADWV 233
 QY 63 QPVTGISQRYRMNSKEKDKLTGLGVLTVMVYIIAIGAGIIGVYKRGKDLKECHDQ 122
 Db 234 TTVWGSSQPVVTR-----GTTDNLIPV-YCSILAAVVGVLAAYIAFKRMNSCK-QNKQ 284
 QY 123 KVCREEMQRITLPISATNPCEIVDEKIVVHTSQPVDPQSGSTPLMGQA 174
 Db 285 GANSR-----PVNQTPPEGEKIHSDSGISVDSQSLHDQHTQTASGQA 329
 RESULT 7
 ID GYG2 HUMAN STANDARD; PRT; 501 AA.
 AC 015486; 015489; 015490; 015486; 015485; 015487;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Glycogenin-2 (EC 2.4.1.186) (GN2).
 GN GYG2.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RC TISSUE=Liver;
 RX MEDLINE=98010589; PubMed=9346895;
 RA Mu J., Skurat A.V., Roach P.J.;
 RT "Glycogenin-2, a novel self-glucosylating protein involved in liver
 glycogen biosynthesis.";
 RL J. Biol. Chem. 272:27589-27597(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20184741; PubMed=10721716;
 RA Zhai L., Mu J., Zong H., DePaoli-Roach A.A., Roach P.J.;
 RT "Structure and chromosomal localization of the human glycogenin-2 gene
 GYG2.";
 RL Gene 242:229-235(2000).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM BETA).
 RC TISSUE=Skin;
 RA Strausberg R.;
 RN Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RP CHARACTERIZATION, AND CARBOHYDRATE-LINKAGE SITE TYR-228.
 RX MEDLINE=99074257; PubMed=9857012;
 RA Mu J., Roach P.J.;
 RT "Characterization of human glycogenin-2, a self-glucosylating
 initiator of liver glycogen metabolism.";
 RL J. Biol. Chem. 273:34850-34856(1998).
 RN [5]
 RP MASS SPECTROMETRY.
 RC TISSUE=Breast cancer;

RX		MEDLINE=21829512; PubMed=11840567;	
RA	Harris R.A., Yang A., Stein R.C., Lucy K., Brusten L., Hereth A., Parekh R., Waterfield M.D., O'Hare M.J., Neville M.A., Page M.J., Zvelebil M.J.;	"Cluster analysis of an extensive human breast cancer cell line protein expression map database."	
RT		Proteomics 2:212-223(2002).	
RL		FUNCTIO: SELF-GLYCOSYLATES, VIA AN INTER-SUBUNIT MECHANISM, TO FORM AN OLIGOSACCHARIDE PRIMER THAT SERVES AS SUBSTRATE FOR GLYCOGEN SYNTHASE.	
CC	-I-	CATALYTIC ACTIVITY: UDP-glucose + glycogenin = UDP + glucosylglycogenin.	
CC	-I-	COCATOR: SELF-GLYCOSYLATION IS DEPENDENT ON THE PRESENCE OF DIVALENT METAL IONS OF WHICH MANGANESE ION IS THE MOST EFFECTIVE.	
CC	-I-	PATHWAY: GLYCOGEN BIOSYNTHESIS.	
CC	-I-	SUBUNIT: HOMODIMER, TIGHTLY COMPLEXED TO GLYCOGEN SYNTHASE.	
CC	-I-	ALTERNATIVE PRODUCTS: AT LEAST 6 ISOFORMS; ALPHA (SHOWN HERE), BETA, GAMMA, DELTA, EPSILON AND ZETA, ARE PRODUCED BY ALTERNATIVE SPLICING.	
CC	-I-	TISSUE SPECIFICITY: EXPRESSED PREFERENTIALLY IN LIVER, HEART, AND PANCREAS.	
CC	-I-	PMT: SELF-GLYCOSYLATED BY THE TRANSFER OF GLUCAN RESIDUES FROM UDP-GLUCOSE TO ITSELF, FORMING AN ALPHA-1,4-GLUCAN OF AROUND 10 RESIDUES ATTACHED TO TYR-228.	
CC	-I-	MASS SPECTROMETRY: MW=55211.89; METHOD=MALDI.	
CC	-I-	SIMILARITY: BELONGS TO THE GLYCOGENIN FAMILY.	
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).	
DR	EMBL; U94362; AAB84377.1; -		
DR	EMBL; U94363; AAB84378.1; -		
DR	EMBL; U94364; AAB84379.1; -		
DR	EMBL; U94357; AAB84373.1; -		
DR	EMBL; U94358; AAB84374.1; -		
DR	EMBL; U94360; AAB84375.1; -		
DR	EMBL; U94361; AAB84376.1; -		
DR	EMBL; AF179624; AAP61855.1; -		
DR	EMBL; AF179615; AAP61855.1; JOINED.		
DR	EMBL; AF179616; AAP61855.1; JOINED.		
DR	EMBL; AF179617; AAP61855.1; JOINED.		
DR	EMBL; AF179618; AAP61855.1; JOINED.		
DR	EMBL; AF179619; AAP61855.1; JOINED.		
DR	EMBL; AF179620; AAP61855.1; JOINED.		
DR	EMBL; AF179621; AAP61855.1; JOINED.		
DR	EMBL; AF179622; AAP61855.1; JOINED.		
DR	EMBL; AF179623; AAP61855.1; JOINED.		
DR	EMBL; BC023152; AAH23152.1; -		
DR	Genev; HGNC:4700; GYG2.		
DR	MIM; 300198; -		
DR	InterPro: IPRO02495; GT 8.		
DR	Pfam; PF01501; Glyco_transf_8; 1.		
KW	Transferase; Glycogen biosynthesis; Alternative splicing;		
KM	Glycoprotein; Polymorphism.		
FT	ACT SITE	119	
FT	CARBOHYD	228	
FT	VARSPLC	3	O-LINKED (GLC..).
FT	VARSPLC	33	MISSING (IN ISOFORM BETA).
FT	VARSPLC	42	MISSING (IN ISOFORM GAMMA).
FT	VARSPLC	378	MISSING (IN ISOFORM DELTA).
FT	VARSPLC	407	MISSING (IN ISOFORM EPSILON).
FT	VARSPLC	413	MISSING (IN ISOFORM ZETA).
FT	VARIANT	270	V->A.
FT	MUTAGEN	228	/FtId=VAR_010401.
FT	MUTAGEN	230	Y->F: LOSS OF ACTIVITY.
FT	CONFLICT	313	Y->P: NO LOSS OF ACTIVITY.
FT	CONFLICT	413	H->R (IN REF. 3).
FT	CONFLICT	462	MISSING (IN REF. 1); AAB84378).
FT	CONFLICT	464	EKV->AGI (IN REF. 1); AAB84376).

[illegible]

Matches 35; Conservative 17; Mismatches 69; Indels 8; Gaps 4;

QY 8 EKRCEDLRCEFTSQALPFTTIOASRGPADEVQVAPANA-LIPARSAAVQPV- 65
 Db 250 EMESRRYSSEAYQALHSL-SSGASFPALPMETVRAPSPNEFLVTRSTKAEVTV 306
 QY 66 -IGISORVNMNSKEKDLGTLGYVLGTMVVIITAGAGIIGVSKRKDKJEQHDQK 123
 Db 307 KPVGNSHNNNSNNNGSKINTLTLVL-IGIIVTAGLGGSFITTOQKEAARAQAERE 365
 QY 124 VCEREMQRI 132
 Db 366 KOEAEQKRI 374

RESULT 9
 TNFR8_HUMAN STANDARD; PRT; 595 AA.
 AC P28908;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tumor necrosis factor receptor superfamily member 8 precursor (CD30L receptor) [lymphocyte activation antigen CD30] (KI-1 antigen).
 GN TNFRSF8 OR CD30.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A. (LONG ISOFORM).
 RC TISSUE=Lymphoid;
 RX MEDLINE=92154659; PubMed=1310894;
 RA Duerkop H., Latza U., Hummel M., Bielebach P., Seed B., Stein H.;
 RT "Molecular cloning and expression of a new member of the nerve growth factor receptor family that is characteristic for Hodgkin's disease";
 RT Cell 68:421-427(1992).
 RL [2]
 RN SEQUENCE FROM N.A. (LONG ISOFORM).
 RP MEDLINE=95089787; PubMed=7527901;
 RA Jung W., Krueger S., Renner C., Gause A., Sahin U., Trumper L.,
 RA Pfeundschuh M.;
 RT "Opposite effects of the CD30 ligand are not due to CD30 mutations: results from cDNA cloning and sequence comparison of the CD30 antigen from different sources";
 RT Mol. Immunol. 31:1329-1334(1994).
 RL [3]
 RN SEQUENCE FROM N.A. (SHORT ISOFORM).
 RP MEDLINE=96437016; PubMed=8839832;
 RA Horie R., Ito K., Tatewaki M., Nagai M., Aizawa S.,
 RA Higashihara M., Ishida T., Inoue J., Takizawa H., Watanabe T.;
 RT "A variant CD30 protein lacking extracellular and transmembrane domain is induced in HL-60 by tetradecanoylphorbol acetate and is expressed in alveolar macrophages";
 RT Blood 88:2422-2432(1996).
 RL [4]
 RN -FUNCTION: Receptor for TNFRSF8/CD30L. May play a role in the regulation of cellular growth and transformation of activated lymphoblasts. Regulates gene expression through activation of NF-kappaB.
 CC -SUBCELLULAR LOCATION: Type I membrane protein (long isoform); Cytoplasmic (short isoform).
 CC -ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a short form/cytoplasmic/CD30V; are produced by alternative initiation.
 CC -PTM: PHOSPHORYLATED ON SERINE AND TYROSINE RESIDUES.
 CC -DISEASE: MOST SPECIFIC HODGKIN'S DISEASE ASSOCIATED ANTIGEN.
 CC -SIMILARITY: CONTAINS 6 TNFR-CYS REPEATS.
 CC -1- DATABASE: NAME=PRO; NOTE=CD guide CD30 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/ncbi/pro/cd/cd30.htm".
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announce/> or send an email to license@ebi.ac.uk).

CC EMBL; M83554; AAA51947.1; -
 CC EMBL; S75768; AAD14188.1; -
 CC EMBL; D86042; BAA12973.1; -
 CC PIR; A42086; A42086.
 CC GeneW; HGNC:11923; TNFRSF8.
 CC MIM; 153243; -
 DR InterPro; IPR001368; TNFR_c6.
 DR Pfam; PF00020; TNFR_c6; 4.
 DR SMART; SM00208; TNFR; 4.
 DR PROSITE; PS00652; TNFR_NGFR_1; 2.
 DR PROSITE; PS50050; TNFR_NGFR_2; 2.
 KW Receptor; Transmembrane; Glycoprotein; Signal; Repeat; Phosphorylation; Alternative initiation.
 FT SIGNAL 1
 FT CHAIN 19 595
 FT FT
 FT CHAIN 464 595
 FT FT
 FT INT MET 464 464
 FT DOMAIN 19 379
 FT TRANSMEM 380 407
 FT DOMAIN 408 595
 FT REPEAT 28 66
 FT REPEAT 68 106
 FT REPEAT 107 150
 FT REPEAT 205 241
 FT REPEAT 243 281
 FT REPEAT 282 325
 FT DOMAIN 347 377
 FT DISULFID 29 44
 FT DISULFID 45 58
 FT DISULFID 48 65
 FT DISULFID 69 81
 FT DISULFID 84 98
 FT DISULFID 87 106
 FT DISULFID 108 122
 FT DISULFID 131 149
 FT DISULFID 233 240
 FT DISULFID 244 256
 FT DISULFID 259 273
 FT DISULFID 262 281
 FT DISULFID 283 297
 FT DISULFID 289 300
 FT CARBOHYD 101 101
 FT CARBOHYD 276 276
 SQ SEQUENCE 595 AA; 63747 MW; 7A407CC78A6E0B08 CRC64;

Query Match 8.2%; Score 75; DB 1; Length 595;
 Best Local Similarity 23.1%; Pred. No. 23;
 Matches 43; Conservative 31; Mismatches 70; Indels 42; Gaps 10;

QY 13 EKRCETTSQALPAT-----TEIQASRGPADEVQVAPANALIPARSAAVQPVIGI 68
 Db 315 QDMAKDTTFEAPLPDGDPCNPTENGEPAST-----SPQSLVLSQASKTLP-IPT 368
 QY 69 SQFRANNSKEKDLGTLGYVLGTMVVIITAGAGIIGVSKRKDKJEQHDQVCERE 128
 Db 369 SAAVVALSTGCKPVLDL-GVYLFWIIVLVVVVSSAFL-----LCHRRACKRR 415
 QY 129 MGR---ITLPLAFTNPICEIVDEKTVVHTSQT-----PVDDQEG--STPLMGQ 173
 Db 416 IRQKHLCPYVOT-SQPKLELVDSRP---RRSSTQLRSGASVTEVAERGLMGQPLMET 471
 QY 174 AGPGGA 179
 Db 472 CHSVGA 477

```
RESULT 10
YFGF_ECOLI STANDARD; PRT; 747 AA.
ID YFGF_ECOLI
AC P77172;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein yfgf.
OS YFGF OR B2503.
OC Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
GN NCBI_TaxID=562;
RX MEDLINE=97426617; PubMed=9278503;
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RL "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
[2]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RA MEDLINE=97426617; PubMed=9278503;
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RL "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
[2]
RP SEQUENCE FROM N.A.
RX STRAIN=K12;
RA MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Alba H., Baba T., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nishimoto H.,
RA Ohshima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasubramam S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horinouchi T.;
RL "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT K-12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features.";
RL DNA Res. 4:91-113(1997).
CC -1 SIMILARITY: CONTAINS 1 DUP2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
DR EMBL; AE000336; AAC75556.1; -.
DR EMBL; D90880; BAA16393.1; -.
DR EcoGene; EG14202; YFGF.
DR InterPro; IPR001633; EAL.
DR InterPro; IPR000160; GGDFF.
DR Pfam; PF00563; EAL; 1.
DR SMART; SM00267; DUF1; 1.
DR SMART; SM00052; DUF2; 1.
KW Hypothetical protein; Complete proteome.
FT DOMAIN 492 736 DUF2.
SQ SEQUENCE 747 AA; 85607 MW; 7D5EBB0BE46C8EDF CRC64;

Query Match 8.1%; Score 74.5; DB 1; Length 747;
Best Local Similarity 25.2%; Pred. No. 33;
Matches 31; Conservative 15; Mismatches 52; Indels 25; Gaps 3;

QY 45 QVFAANLIPARSEAAVQPIVIGISQRNMSKEKDGTL-----GVVLTITMMVI 96
DB 141 QPAAVVGLASRENLVGVMP-----NLGTILNQALIVGNLIGVPLCYF 185
QY 97 IIAIGAGIILGYSYRGKDLKEQHDQYCEBEMQRTITPLSAFTNPCEIYDEKTVVHT 156
DB 166 IIRVRNPFYLRSY--SOLKQVDAAKYTKKEFALMLALGALLLLCMPLNEKSTIFST 243
QY 157 SQT 159
```

```
Db 244 NYT 246

RESULT 11
MTXB_METEX STANDARD; PRT; 260 AA.
ID MTXB_METEX
AC P53595;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Malate--CoA ligase alpha chain (EC 6.2.1.9) (Malyl-CoA synthetase)
DE (Malate thiokinase) (MTX-alpha).
OS MTXB.
OS Methylotuberculosis extorquens.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Methylotubacterium group; Methylotubacterium.
GN NCBI_TaxID=408;
RX MEDLINE=95050329; PubMed=7961516;
RA STRAIN=AM1 / NCIMB 9133;
RA MEDLINE=95050329; PubMed=7961516;
RA Chistoserdova L.V., Lidstrom M.E.;
RL "Genetics of the serine cycle in Methylotubacterium extorquens AM1:
RT identification, sequence, and mutation of three new genes involved in
RT C1 assimilation, orf4, mtbA, and mtbB.";
RL J. Bacteriol. 176:7398-7404(1994).
CC -1 CATALYTIC ACTIVITY: ATP + malate + CoA = ADP + phosphate + malyl-
CC CoA.
CC -1 PATHWAY: SERINE CYCLE.
CC -1 SUBUNIT: COMPOSED OF AN ALPHA CHAIN AND A BETA CHAIN.
CC -1 SIMILARITY: TO OTHER ALPHA SUBUNITS OF SUCCINYL-CoA SYNTHETASE,
CC OF MALATE--COA LIGASE AND TO ATP CITRATE-LYASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
DR EMBL; L33465; AAA62655.1; -.
DR HSP; P07459; 28CU.
DR InterPro; IPR003781; CoA binding.
DR InterPro; IPR000303; CoA ligase.
DR Pfam; PF00549; ligase-CoA; 1.
DR Pfam; PF02829; CoA binding; 1.
DR PROSITE; PS01216; SUCCINYL-CoA_LIG_1; 1.
DR PROSITE; PS00399; SUCCINYL-CoA_LIG_2; 1.
KW Ligase; Phosphorylation; ATP-binding.
FT ACT_SITE 251 251
FT CATALYSTS (BY SIMILARITY).
SQ SEQUENCE 260 AA; 26978 MW; 9A26A3B057A4A447 CRC64;

Query Match 8.0%; Score 73; DB 1; Length 260;
Best Local Similarity 22.6%; Pred. No. 14;
Matches 43; Conservative 30; Mismatches 65; Indels 52; Gaps 8;

QY 21 TSOALPAFTTIOEASEGGA-DEQVAPAPNALPARESEA-----AYQPV 65
DB 47 THGVGVFNT-VKEAVEATGATSTTPVAPPPAADLMEADAGLKVCSITDGIAPADM 105
QY 66 IGISQVRNMSKEKDLGTLGYVLGITMMV-----IIAIG--AGIILGYSYRGKDLK 117
DB 106 MRYKRYLRIRYPRKEK-----TWVGNVCAGIISPGKSMGIMGHYLYPGK--- 151
QY 118 EQHDQYCE-----EKQRTITPLSAFTNPCEIYDEKTVVHTSQTPVDPQEGS 167
DB 152 ---GVISRGTLGYEAQAQMKELGISTSVGIGDPINGSPLDHALPEQDPTEA 207
QY 168 TPLMGQAGTP 177
```

```
Db 208 VLMIGEIGP 217

RESULT 12
IDENT_SCHPO STANDARD; PRT; 536 AA.
AC 013695;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical serine-rich protein C11G7.01 in chromosome I.
GN SPAC11G7.01.
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetaceae; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
NCBI_TaxID=4896;
RN NCB1
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Soutos J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsle K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean U.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitz E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Wolcraet G., Aert R., Robben J., Grymoprez B.,
RA Welfjens I., Vanstreels E., Rieger W., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs W., Fritz C., Holzer E., Mosel D., Hilbert H.,
RA Gaborz K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motier S.,
RA Galbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Umenetz J., Sanchez S., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC CC
DR EMBL; 299161; CAB16206.1; -
KW Hypothetical protein; Transmembrane.
FT DOMAIN 25 238 SRR-RICH.
FT TRANSMEM 247 257 POTENTIAL.
FT TRANSMEM 351 371 POTENTIAL.
SQ SEQUENCE 536 AA; 52992 MW; CA7EFAF19ED658CA CRC64;

Query Match 8.0%; Score 73; DB 1; Length 536;
Best Local Similarity 36.4%; Pred. No. 31;
Matches 20; Conservative 10; Mismatches 21; Indels 4; Gaps 1;

Qy 76 SKRKDLGLGLVYLG---ITMWTIIAIGAGIILGYSYKRGKDKKEQHDQKVC 126
Db 232 SKSNKSNHVGVGCSVAIPGVVLIILGLGIFLMKRRQKRIKAEKMEVEE 286

RESULT 13

ICP0_HSVBU
ID ICP0_HSVBU STANDARD; PRT; 676 AA.
AC P29128;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Trans-acting transcriptional protein ICP0 (p135 protein) (IER
DE 2.9/ER2.6).
GN BICP0.
OS Bovine herpesvirus type 1 (strain Jura).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicelloviruses.
NCBI_TaxID=31518;
RN NCB1
RP SEQUENCE FROM N.A.
RX MEDLINE=92219360; PubMed=1313901;
RT Wirth U.V., Fraefel C., Vogt B., Vicek C., Paces V., Schwyzler M.;
RT "Immediate-early RNA 2.9 and early RNA 2.6 of bovine herpesvirus 1
RT are 3' coterminal and encode a putative zinc finger transactivator
RT protein.";
RL J. Virol. 66:2763-2772(1992).
CC CC
CC -1- PTM: THE STRONGLY ACIDIC REGION MIGHT SERVE AS A TRANSCRIPTIONAL
CC ACTIVATION DOMAIN, POSSIBLY REGULATED THROUGH PHOSPHORYLATION BY
CC CASEIN KINASE II.
CC CC
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC CC
DR EMBL; M84465; AAA46062.1; -
DR PIR; A38209; EDRE23.
DR HSP; P28990; ICHC.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00097; zf-C3HC4_1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PSS0089; ZF_RING_2; 1.
KW Transcription regulation; Trans-acting factor; Activator; Zinc-finger;
KW DNA-binding; Early protein; Repressor; Phosphorylation.
FT ZN_FING 13 52 RING-TYPE.
FT DOMAIN 284 331 ASF/GLU-RICH (ACIDIC).
SQ SEQUENCE 676 AA; 67879 MW; 11B06BA4E5C4EB71 CRC64;

Query Match 8.0%; Score 73; DB 1; Length 676;
Best Local Similarity 23.2%; Pred. No. 40;
Matches 47; Conservative 21; Mismatches 95; Indels 40; Gaps 8;

Qy 2 GEAVPEKRCPCEDIRCPETTSOALPFTTBIOGASGPGADVQVAPANA----- 52
Db 336 GASGAARARPRARQYVSTGRTPAVQAPRSLARPCRAAAVAPSSRRGGRDP 395
Qy 53 -LPRSSAAAVPVYIGISQVRNMSKSKDGLG---LGYVLGITMWTIIAIGAGIILGY 108
Db 396 RLPAAPAPAA-----AAQARACSPPEPRGAGGVAAGET-----AGWGAGSEEG 443
Qy 109 SYKRGDKLKEQHOKYCR-----EMQRTLPISA---FTNPCEIVDEKTVVHTSOTP 160
Db 444 GERPARLIGAGPRVQARRRRRTTELDRAPTAPAPAPAPAPAPAPAPAPAPAPAP 500
Qy 161 VDPQEGSTP-----LMGQAGTPGA 179
Db 501 ADPAPAPAPAPAPAGAGQIGTPAA 523

RESULT 14
EFG_DEIRA
```

```

ID   BEG_DEIRA          STANDARD;          PRT;          698 AA.
AC   O9RKS;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Elongation factor G (EF-G).
GN   FUSA OR DR0307.
OS   Deinococcus radiodurans.
OC   Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC   Deinococcaceae; Deinococcus.
OX   NCBI_TaxID=1299;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R1.
RX   MEDLINE=20036896; PubMed=10567266;
RA   Dodson R.J., Haft J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA   Wofat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA   Vamathevan J., Lam P., McDonald L., Utecherback T., Zaleski C.,
RA   Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA   Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA   Fraser C.M.;
RT   "Sequence of the radioresistant bacterium Deinococcus
RT   radiodurans R1."
CC   Science 286:1571-1577(1999).
CC   -1- FUNCTION: This protein promotes the GTP-dependent translocation of
CC   the nascent protein chain from the A-site to the P-site of the
CC   ribosome.
CC   -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC   EF-G/EF-2 SUBFAMILY.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; AE001891; AAF09887.1; -.
DR   HSSP; P13551; 1FNM.
DR   TIGR; DR0307; -.
DR   InterPro; IPR0004540; EF-G.
DR   InterPro; IPR0004640; EF-G.
DR   InterPro; IPR004161; EFU_D2.
DR   InterPro; IPR000795; EF_GTPbind.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00679; EFG_C_1.
DR   Pfam; PF00009; GTP_EFTU_1.
DR   Pfam; PF03144; GTP_EFTU_D2_1.
DR   PRINTS; PR00315; BLONGATNFC.
DR   TIGRPFAM; TIGR00484; EF-G; 1.
DR   TIGRPFAM; TIGR00231; small_GTP_1.
DR   PROSITE; PS00301; EFACOR_GTP_1.
KW   Elongation factor; Protein biosynthesis; GTP-binding;
KM   Complete proteome.
FT   NP_BIND          20      GTP (BY SIMILARITY).
FT   NP_BIND          90      GTP (BY SIMILARITY).
FT   NP_BIND          144     GTP (BY SIMILARITY).
FT   NP_BIND          147     GTP (BY SIMILARITY).
SQ   SEQUENCE 698 AA; 76798 MW; 92E03D5645B5F49B CRC64;

Query Match          8.0%; Score 73; DB 1; Length 698;
Best Local Similarity 24.8%; Pred. NO. 42; Indels 16; Gaps 5;
Matches 40; Conservative 22; Mismatches 81;

OY   26 PAFTTEIQEAS-----EGPQADEVOVFAPANALPARSFAAIVOPUG-----ISQRYRMS 76
DB   444 PTFEVEEDDESQGTITAGGEHLHLIVDRLRREYKVAVANGAQVAVYRETIKQVEVDS 503
OY   77 KEKKDLGTLGVVLTIMVYIIAIGAGIILGYSGKGLKEQ--HQQKVCEREMRITL 134
DB   504 KFNARQSGRGQYGVHVLKRLREPLEPGAGIFENAVVGVGVPEYVIGPRAQKVEVAMQ--SG 561

```

```

OY   135 PLSAFTNPTCEIVDEKTVVHTSOTRPDPQSGTPLYMCQAG 175
DB   562 PMLGF-----PVYDIKVIYIDSGSYHEVDSSEMAFKIASMG 597

RESULT 15
ID   B3AR_MOUSE        STANDARD;          PRT;          400 AA.
AC   P25962; O9QZ98;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Beta-3 adrenergic receptor.
GN   ADRB3 OR ADRB3R OR B3BAR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Swiss;
RX   MEDLINE=92037534; PubMed=1718744;
RA   Nahmias C., Blin N., Elalouf J.M., Mattei M.-G., Strosberg A.D.,
RA   Emorine L.J.;
RT   "Molecular characterization of the mouse beta3-adrenergic receptor:
RT   relationship with the atypical receptor of adipocytes."
RT   EMBO J. 10:3721-3727(1991).
RN   [2]
RP   REVISIONS, SEQUENCE FROM N.A.
RX   MEDLINE=93279311; PubMed=8389293;
RA   van Spronsen A., Nahmias C., Krief S., Briand-Sutren M.-M.,
RA   Strosberg A.D., Emorine L.J.;
RT   "The promoter and intron/exon structure of the human and mouse beta
RT   3-adrenergic-receptor genes."
RT   Eur. J. Biochem. 213:1117-1124(1993).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM B).
RC   STRAIN=C57BL/6; TISSUE=Brown adipose tissue;
RX   MEDLINE=9936494; PubMed=1045305;
RA   Evans B.A., Papaioannou M., Hamilton S., Summers R.J.;
RT   "Alternative splicing generates two isoforms of the beta3-adrenoceptor
RT   which are differentially expressed in mouse tissues."
RT   Br. J. Pharmacol. 127:1525-1531(1999).
RN   [4]
RP   SEQUENCE OF 378-400 FROM N.A.
RC   TISSUE=Adipose tissue;
RX   MEDLINE=93125503; PubMed=1336117;
RA   Gramanman J.G., Labners K.N., Rao D.D.;
RT   "Rodent and human beta 3-adrenergic receptor genes contain an intron
RT   within the protein-coding block."
RT   Mol. Pharmacol. 42:964-970(1992).
CC   -1- FUNCTION: BETA-ADRENERGIC RECEPTORS MEDIATE THE CATECHOLAMINE-
CC   INDUCED ACTIVATION OF ADENYLATE CYCLASE THROUGH THE ACTION OF G
CC   PROTEINS. BETA-3 IS INVOLVED IN THE REGULATION OF LIPOLYSIS AND
CC   THERMOGENESIS.
CC   -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A (SHOWN HERE) AND B; ARE
CC   PRODUCED BY ALTERNATIVE SPLICING.
CC   -1- TISSUE SPECIFICITY: WHITE AND BROWN ADIPOSE TISSUES, AND DIGESTIVE
CC   TRACT. ISOFORM B HIGHEST IN BRAIN.
CC   -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; X72862; CAAS1384.1; -.
DR   EMBL; X60438; CAA42966.1; -.

```

```

DR EMBL: AF193027; AAF05768.1; -.
DR EMBL: SS3290; AAB24836.1; -.
DR PIR: S18548; S18548.
DR PIR: S12804; S12804.
DR MGI: MGI:87939; AATB3.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Multigene family; Phosphorylation; Lipoprotein; Palmitate;
KW Alternative splicing
FT DOMAIN 1 36 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 37 60 1 (POTENTIAL).
FT DOMAIN 61 69 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 70 88 2 (POTENTIAL).
FT DOMAIN 89 108 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 109 130 3 (POTENTIAL).
FT DOMAIN 131 152 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 153 175 4 (POTENTIAL).
FT DOMAIN 176 200 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 201 222 5 (POTENTIAL).
FT DOMAIN 223 269 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 290 311 6 (POTENTIAL).
FT DOMAIN 312 323 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 324 344 7 (POTENTIAL).
FT DOMAIN 345 400 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 8 8 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 26 26 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 107 166 BY SIMILARITY.
FT LIPID 358 358 PALMITATE (BY SIMILARITY).
FT VARSPLIC 388 400 REDGTGGRAPPT -> SLLRFRRLHYTCLGTP (IN
SQ SEQUENCE 400 AA; 43006 MW; 374E072C01DFA32E CRC64;
ISOFORM B).

Query Match 7.9%; Score 72.5; DB 1; Length 400;
Best Local Similarity 19.5%; Pred. No. 25;
Matches 39; Conservative 21; Mismatches 57; Indels 83; Gaps 8;

QY 7 PEKPCEDLCPETTSQALPAFTTEIQEASBEGADDEVQFAPANALPARSEMAAVQPI 66
DB 244 PESSPPSPSPSPSPATGTPA-----APDGVPPCGRRPARLPL- 282
QY 67 GISQRYMNSKEKKDLGTLGYLGITM-----VIIAIGAIL-----IG 107
DB 283 -----REHRLATLTGLIMGIFSLCLPFFLANVLRALAGPSLVPSGVFIALNMLG 332
QY 108 YS-----YKRGKDLKEQHDKVCE-----REMORTLPLSAFTNPTCEIVDEKTV 153
DB 333 YANSAFNPVLYCRSPDRDAFRLLCSYGGRGPEPRAVTPFASP----- 377
QY 154 VHTSQTP---VDPQEGSTP 169
DB 378 VEARQSPPLNRFDEGEGARP 397

```

Search completed: April 7, 2003, 09:17:01
 Job time : 12.1871 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:25 ; Search time 48.0062 Seconds
(without alignments)
768.284 Million cell updates/sec

Title: US-10-057-951-2_COPY_85_263

Sequence: 1 SGEAGVPEKPCEDLRCPET.....PVDPGSTPLMGQAGTGA 179

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SEPREMBL_21.*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mtc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.ivirus:*
16: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	916	100.0	263	4 Q96FE7	Q96FE7 homo sapien
2	913	99.7	263	4 Q00318	Q00318 homo sapien
3	913	99.7	263	4 Q00318	Q00318 homo sapien
4	93.5	10.2	569	10 Q9VM61	Q9VM61 drosophila
5	87.5	9.6	759	5 Q20429	Q20429 caenorhabdi
6	86	9.4	290	5 Q9BL62	Q9BL62 caenorhabdi
7	86	9.4	1165	4 Q9C0G0	Q9C0G0 homo sapien
8	85.5	9.3	1115	11 Q91V14	Q91V14 mus musculu
9	85.5	9.3	1116	11 Q63633	Q63633 rattus norv
10	84.5	9.2	607	17 Q8U3P9	Q8U3P9 pyrococcus
11	83.5	9.1	517	16 Q93J57	Q93J57 streptomyce
12	83.5	9.1	733	16 Q9PA31	Q9PA31 xyloella fas
13	83	9.1	1553	5 Q95077	Q95077 caenorhabdi
14	82	9.0	774	10 Q9L128	Q9L128 oryza sativ
15	81.5	8.9	535	10 Q8W4V9	Q8W4V9 solanum tub
16	81.5	8.9	715	10 Q9M839	Q9M839 arabidopsis

17	81.5	8.9	1365	5 Q9NEV4	Q9NEV4 caenorhabdi
18	80.5	8.8	803	16 Q25153	Q25153 heliobacte
19	80.5	8.8	918	16 Q8XR11	Q8XR11 raietsonia s
20	80.5	8.8	1852	3 Q9C2H4	Q9C2H4 neurospora
21	80.5	8.8	2358	16 Q9L1V8	Q9L1V8 streptomyce
22	80	8.7	618	2 Q9F5E5	Q9F5E5 agrobacteri
23	79.5	8.7	770	2 Q9AQ89	Q9AQ89 bacillus th
24	79	8.6	960	16 Q92KFS	Q92KFS rhizobium m
25	78.5	8.6	457	16 Q8UHM4	Q8UHM4 agrobacteri
26	78.5	8.6	478	5 Q25684	Q25684 plasmodium
27	78.5	8.6	552	5 Q9VE62	Q9VE62 drosophila
28	78.5	8.6	2972	5 P90891	P90891 caenorhabdi
29	78	8.5	224	16 Q8Z5M1	Q8Z5M1 salmoneilla
30	78	8.5	336	17 Q8ZDF0	Q8ZDF0 pyrobaculum
31	78	8.5	352	16 Q98BL5	Q98BL5 rhizobium l
32	77.5	8.5	267	16 Q8UB02	Q8UB02 agrobacteri
33	77.5	8.5	947	4 Q8TDV8	Q8TDV8 homo sapien
34	77.5	8.5	947	4 Q8TBZ2	Q8TBZ2 homo sapien
35	77.5	8.5	1230	5 Q20626	Q20626 caenorhabdi
36	77	8.4	624	11 Q70236	Q70236 mus musculu
37	76.5	8.4	518	3 Q9F732	Q9F732 neurospora
38	76.5	8.4	766	2 Q9ZHF3	Q9ZHF3 neisseria m
39	76.5	8.4	1101	4 Q9ULP4	Q9ULP4 homo sapien
40	76.5	8.4	1116	4 Q9H2X9	Q9H2X9 homo sapien
41	76.5	8.4	1116	4 Q9H4Z0	Q9H4Z0 homo sapien
42	76	8.3	453	16 Q8UUV0	Q8UUV0 agrobacteri
43	76	8.3	482	5 Q93998	Q93998 plasmodium
44	76	8.3	847	6 Q95KP7	Q95KP7 canis fami1
45	76	8.3	1406	5 Q01777	Q01777 caenorhabdi

ALIGNMENTS

RESULT 1

Q96FE7 PRELIMINARY; PRT; 263 AA.
AC Q96FE7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Unknown (protein for MG:17330).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011049; AA011049.1; -
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR ProDom; PD000395; Kringle; 1.
DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_1.
DR PROSITE; PS50070; KRINGLE_2; 1.
SQ SEQUENCE 263 AA; 28234 MW; 197C3BEE888FA242 CRC64;

Query Match 100.0%; Score 916; DB 4; Length 263;
Best Local Similarity 100.0%; Pred. No. 6.6e-78;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	SGEAGVPEKPCEDLRCPETTSQALPAFTTEIQEASBPGADDEVQVAPNALPARESEA	60
DB	85	SGEAGVPEKPCEDLRCPETTSQALPAFTTEIQEASBPGADDEVQVAPNALPARESEA	144
QY	61	AVQPVIGISQVSMSEKKKDLGTIGYVIGITMWVITAIAGAGIILGYSYKRGKDLSEQH	120
DB	145	AVQPVIGISQVSMSEKKKDLGTIGYVIGITMWVITAIAGAGIILGYSYKRGKDLSEQH	204
QY	121	DQKCEKEMQRIPLPSATFNPCETVDEKTVVHTSQTVPDPEGSTPLMGQAGTGA	179

```

Db 205 DQKVCEREMQRIITLPLSAFTNPCEIVDEKTVVHTSQTPVDQEGSTPLMGQAGTPGA 263
RESULT 2
ID 000318 PRELIMINARY; PRT; 263 AA.
AC 000318;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE MGSC:DJ515N1.2 protein.
GN MGSC:DJ515N1.2.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Du Z., Scheet P., Harper M.,
RT "The sequence of H. sapiens PAC clone RP3-S15N1."
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC002073; AAB54054.1; -.
DR HSSP; P00749; IKDU.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE; FALSE_NEG.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
SQ SEQUENCE 263 AA; 28248 MW; 1973EEB854A242 CRC64;

Query Match 99.7%; Score 913; DB 4; Length 263;
Best Local Similarity 99.4%; Pred. No. 1.3e-77;
Matches 178; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SGEAGVPEKRCEDLRCPEITTSQALPAFTTEIQASGPGADVOVPANALPARSEAA 60
Db 85 SGEAGVPEKRCEDLRCPEITTSQALPAFTTEIQASGPGADVOVPANALPARSEAA 144
Qy 61 AVQPVIGISQVRNMSKEKDLGTLGVLTGMVVIITAGAGIIGYKRGKDLKEQH 120
Db 145 AVQPVIGISQVRNMSKEKDLGTLGVLTGMVVIITAGAGIIGYKRGKDLKEQH 204
Qy 121 DQKVCEREMQRIITLPLSAFTNPCEIVDEKTVVHTSQTPVDQEGSTPLMGQAGTPGA 179
Db 205 DQKVCEREMQRIITLPLSAFTNPCEIVDEKTVVHTSQTPVDQEGSTPLMGQAGTPGA 263

RESULT 3
Qy 09V6V4 PRELIMINARY; PRT; 8805 AA.
AC 09V6V4;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Shot protein.
GN SHOT OR CG18076 OR CG18637.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=BERKELEY;
RX MEDLINE=20196006; Pubmed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,

```

```

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortan J.R., Yandell M.D., Zhang O., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blarej R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Adair J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Bendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertan B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,
RA Burtis K.C., Buzan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Giodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kratz C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Paclob J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskis R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195 (2000).
DR EMBL; AE003816; AAG22268.1; -.
DR HSSP; O01082; IBKR.
DR FlyBase; FBgn001373; shot.
DR InterPro; IPR001715; Shc-like.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003108; GAS2.
DR InterPro; IPR000005; HTHAc.
DR InterPro; IPR001101; Plectin_repeat.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR002017; Spectrin.
DR Pfam; PF00307; CH; 1.
DR Pfam; PF00306; ehand; 2.
DR Pfam; PF02187; GAS2; 1.
DR Pfam; PF00018; SH3; 1.
DR Pfam; PF00435; spectrin; 36.
DR ProDom; PD000012; EF-hand; 1.
DR SMART; SM00033; CH; 1.
DR SMART; SM00054; Erb; 2.
DR SMART; SM00243; GAS2; 1.
DR SMART; SM00250; PLEC; 19.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00150; SPEC; 34.
DR PROSITE; PS50021; CH; 1.
DR PROSITE; PS50018; EF_HAND; UNKNOWN_1.
DR PROSITE; PS00041; HTH_ARAC_FAMILY_1; UNKNOWN_1.
SQ SEQUENCE 8805 AA; 989540 MW; 67E4BDCC1DD83F CRC64;

Query Match 10.3%; Score 94; DB 5; Length 8805;
Best Local Similarity 26.3%; Pred. No. 54;
Matches 45; Conservative 19; Mismatches 83; Indels 24; Gaps 8;

Qy 8 EKRPCEDLRCPEITTSQALPAFTTEIQASGPGADVOVPANALPARSEAA 60
Db 3379 EITPTEPISSEITVITLTKTIVDEKGLPPSIQIOPHSISDQGLVDFIEBK 3438
Qy 61 AVQPVIGISQVRNMSKEKDLGTLGVLTGMVVIITAGAGIIGYKRGKDLKEQH 120

```


Db 3439 GTPABEG-GSN10NVKXDAKLGALA-IVGAPVL-----AGKALVDALTEKTLKREH 3489
 QY 121 --DQRYCEREMQRTPLPLSAFTNPCEIVDEKTVVHTSOTPVDPQEGSTP 169
 Db 3490 STSQVTEVEE---LPLEIITTSATLVTKVTTT--TTTTTSTSEG-TP 3534

RESULT 4

Q9M261 PRELIMINARY; PRT; 569 AA.
 AC Q9M261;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE Hypothetical 63.5 kDa protein.
 GN F7M19.10.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eustosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RX NCBI_TaxID=3702;
 RP SEQUENCE FROM N.A.
 RA Nakamura G., Fartmann B., Dauner D., Sterr W., Holland R.,
 RA Weichselgartner M., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.,
 RA Quecter F., Salenouat M.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN 121
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL338643; CAB86471.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 569 AA; 63537 MW; B75942AC7C565F5 CRC64;

Query Match 10.2%; Score 93.5; DB 10; Length 569;
 Best Local Similarity 32.1%; Pred. No. 2.1;
 Matches 27; Conservative 10; Mismatches 36; Indels 11; Gaps 1;

QY 8 EKRPCEDLRCPETTSQALPAFTTEIOASBEGADVEQVAPAN-----ALPAP 56
 Db 301 KKRAERRRVVEGSPAPSGAREVPEVAAPAPVDVDPAPRPETATQAVIALPAR 360
 QY 57 SEAAAVQPVIGISQVRMNSKEKK 80
 Db 361 DKASGKSPQIDTSQEKRRKKKKK 384

RESULT 5

Q20429 PRELIMINARY; PRT; 759 AA.
 AC Q20429;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE F45B12.2 protein.
 GN F45B12.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OC NCBI_TaxID=6239;
 RX NCBI_TaxID=6239;
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kerstew J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Ritken L., Roopra A., Saunders R., Showkeen R.,
 RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,

RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterson R.,
 RA Watson A., Weinstock L., Wilkinson-Sprot J., Wohlman P.; III of C.
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38(1994).
 RN 121
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Latreille P.;
 RL "The sequence of C. elegans cosmid F45B12.";
 RN Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 RN 121
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterson R.;
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U29516; AA68790.1; -.
 DR InterPro; IPR004366; Cyclin.
 DR InterPro; IPR000812; TFIIB euk.
 DR Pfam; PF00382; transcript_fac2; 2.
 DR SMART; SM00385; CYCLIN_2.
 SQ SEQUENCE 759 AA; 84033 MW; 1645547741D3043 CRC64;

Query Match 9.6%; Score 87.5; DB 5; Length 759;
 Best Local Similarity 30.4%; Pred. No. 11;
 Matches 28; Conservative 17; Mismatches 32; Indels 15; Gaps 5;

QY 1 SGEAGVPEKRPEDL--RCPETTSQALPAFTTEIOASBEGADVE-----QVAPANAL 53
 Db 577 ASESTIQKRSIFDLTEBCESEKSNKSPKVLKVSAS--PSTSEVSIHKFPV----- 630
 QY 54 PARSEAAVQPVIGISQVRMNSKEKKDLGTL 85
 Db 631 PARSRVAKKPTTGAKKLAALN--EVKVVHY 660

RESULT 6

Q9BL62 PRELIMINARY; PRT; 290 AA.
 AC Q9BL62;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical 33.3 kDa protein.
 GN Y65B4.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OC NCBI_TaxID=6239;
 RX NCBI_TaxID=6239;
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN 121
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Maupin R.;
 RL "The sequence of C. elegans cosmid Y65B4.";
 RN Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN 121
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX "Direct Submission.";
 RA Waterson R.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC024844; AAK29959.2; -.
 KW Hypothetical protein.
 SQ SEQUENCE 290 AA; 33264 MW; 1BD039FDBA67EA77 CRC64;

```

Query Match          9.4%; Score 86; DB 5; Length 290;
Best Local Similarity 18.8%; Pred. No. 4.7;
Matches 30; Conservative 29; Mismatches 55; Indels 46; Gaps 4;

Qy 7 PERPCEDLRCPPTSQLPAFTTEIQEASSEGPADEVQVFAFANLAPRSEAAVQPVY 66
    |||
    |||
    |||
Db 95 PERK-----KLTDEKETQKVGSHSENYSNFKPLKN-LKLIQELR 137
    |||
    |||
    |||
Qy 67 GISQRYVMSKSKKDLGTLGYVIGITWVVIIGAGIILGYSKGDKLQHQDKVCE 126
    |||
    |||
    |||
Db 138 ENKRRAREAKERKTL-----EKDKKSEMARREKEXKD 172
    |||
    |||
    |||
Qy 127 REMQRTPLPLSAFTNPCEIVDEKTVVHTSQTPVDPQEG 166
    |||
    |||
    |||
Db 173 KETRKIEREISAATNSKCELY---TFCHVQKTVIDTWG 208
    |||
    |||
    |||

RESULT 7
Q9C0G0 PRELIMINARY; PRT; 1165 AA.
ID Q9C0G0
AC Q9C0G0;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE KIAA1703 protein (Fragment).
GN KIAA1703.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082932; PubMed=11214970;
RA Nagaoka T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIX.
RT The complete sequences of 106 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 7:347-355(2000).
DR EMBL; AB051490; BAB21794.1; -.
DR HSSP; P08047; ISP2.
DR InterPro; IPR000822; Znf C2H2.
DR Pfam; PF00096; zf-C2H2_12.
DR SMART; SM00355; Znf C2H2_12.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 5.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 7.
KM DNA-binding; Metal-binding; zinc-finger.
FT NON TER 1
SQ SEQUENCE 1165 AA; 126980 MW; A37B8A9701F5133E CRC64;

Query Match          9.4%; Score 86; DB 4; Length 1165;
Best Local Similarity 24.7%; Pred. No. 26;
Matches 54; Conservative 30; Mismatches 83; Indels 52; Gaps 13;

Qy 2 GEAGVPERK-RP-CEDLRCPETTSQALPAFTTEIQEASRGPADQVQFAFANLAPARSE 58
    |||
    |||
    |||
Db 921 GRAGIEEGRGKADVL-----QLPG--QEVSHVADPFAPIQCFPPAQSEPAVEV 972
    |||
    |||
    |||
Qy 59 -AAAVQPIGISQRYVMSKSKK-----DLGTLGYVLT-GITWVY----- 96
    |||
    |||
    |||
Db 973 LTQVHPSAAMAASQRAQVAFKMWQVGLQPAVCDTAAAGQLVADGVTVQVVSSEGAVM 1032
    |||
    |||
    |||
Qy 97 IIAIGAGIILGYSKGDKLQHQDKVCE---QHQDKVCEMRQRTPLSA--FTNPCTCIV---- 147
    |||
    |||
    |||
Db 1033 VAGSEAGIIMQEAQEHNDLVESDGLISQIITVELVQAMVQESSGGFSGTTHYITEL 1092
    |||
    |||
    |||
Qy 148 -----DEKTVVHTSQTPVDPQEG---GSTPLMGQAGTP 177
    |||
    |||
    |||
Db 1093 PPGVDPPEGLYSHVLETRADSEQLIQAGAT-LGTPAGAP 1130
    |||
    |||
    |||

RESULT 8
Q91V14 PRELIMINARY; PRT; 1115 AA.
ID Q91V14

```

```

AC Q91V14;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE K-C1 cotransporter.
GN K-C12A5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ILS, AND ISS;
RA Ehringer M.A., Thompson J., Conroy O., Xu Y., Yang F., Camniff J.,
RA Beeson M., Gordon L., Bennett B., Johnson T.E., Sikele J.M.;
RT "High-Throughput Sequence Identification of Gene Coding Variants
RT within Alcohol-Related QTLs.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF332064; AAK56093.1; -.
DR EMBL; AF332063; AAK56092.1; -.
DR InterPro; IPR002293; AA/rel_primeasel.
DR InterPro; IPR004842; KCl_cotransport.
DR TrEMBL; TIGR00930; aa30.1.
SQ SEQUENCE 1115 AA; 123587 MW; DD506AC24D8492B4 CRC64;

Query Match          9.3%; Score 85.5; DB 11; Length 1115;
Best Local Similarity 27.1%; Pred. No. 27;
Matches 32; Conservative 15; Mismatches 44; Indels 27; Gaps 4;

Qy 79 KDLGTLGYVIGITWVVIIGAGIILGYSY-----KRGDKLQHQ-DQKVCEREQ 130
    |||
    |||
    |||
Db 872 KDLGTLGYVIGITWVVIIGAGIILGYSY-----KRGDKLQHQ-DQKVCEREQ 931
    |||
    |||
    |||
Qy 131 RIT-----PLSAFTNPTC--EIVDEKTVVHTSQTPVDPQEGSRP 169
    |||
    |||
    |||
Db 932 SITDESRSIRKRPANRLNVPETTAQCNKEKPEEVOLIHQASPCSSPSP 989
    |||
    |||
    |||

RESULT 9
Q63633 PRELIMINARY; PRT; 1116 AA.
ID Q63633
AC Q63633;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Fucose-binding sensitive K-C1 cotransporter.
GN KCC2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE DAWLEY; TISSUE=BRAIN;
RX MEDLINE=96279171; PubMed=8663311;
RA Payne J.A., Stevenson T.J., Donaldson L.F.;
RT "Molecular characterization of a putative K-C1 cotransporter in rat
RT brain. A neuronal-specific isoform.";
RL J. Biol. Chem. 271:16245-16252(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE DAWLEY; TISSUE=BRAIN;
RX MEDLINE=96279170; PubMed=8663127;
RA Gillen C.M., Brill S., Payne J.A., Forbush B. III;
RT "Molecular cloning and functional expression of the K-C1 cotransporter
RT from rabbit, rat, and human. A new member of the cation-chloride
RT cotransporter family.";
RL J. Biol. Chem. 271:16237-16244(1996).
DR EMBL; U55816; AAC52635.1; -.
DR InterPro; IPR002293; AA/rel_primeasel.
DR InterPro; IPR004842; KCl_cotransport.
DR InterPro; IPR000076; KCl_transport.
DR PRINTS; PRO1081; KCLTRANSFPORT.

```

DR TIGR00930; 2a30; 1.
SQ SEQUENCE 1116 AA; 123563 MW; 882FDA0D7FDE2F6 CRC64;
Query Match 9.3%; Score 85.5; DB 11; Length 1116;
Best Local Similarity 27.1%; Pred. No. 27;
Matches 32; Conservative 15; Mismatches 44; Indels 27; Gaps 4;
QY 79 KKDGLTGYVLGITMWTIIAIGAGILGYSY-----KRGDKLKRQH-DQKVCERMQ 130
DB 872 KKDGLTGYVLGITMWTIIAIGAGILGYSY-----KRGDKLKRQH-DQKVCERMQ 130
QY 131 RIT-----LPLSAFTNPTC---EIVDEKTVVHTSQTVPDPOGSGTP 169
DB 932 SITDESGSIRKKNPANTRLANPEETACDNEKPEEFVQLHIDQSPCSPSSPSP 989
RESULT 10
Q8U3P9 PRELIMINARY; PRT; 607 AA.
AC Q8U3P9;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical protein PF0407.
GN PF0407.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VCI / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb P.T., Brown J.R.;
RT "The complete sequence of the EMBL/GenBank/DBJ databases."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE010163; AAL80531.1; -
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 607 AA; 68721 MW; 0D125AC09DE849AD CRC64;
Query Match 9.2%; Score 84.5; DB 17; Length 607;
Best Local Similarity 28.8%; Pred. No. 16;
Matches 36; Conservative 10; Mismatches 38; Indels 41; Gaps 6;
QY 24 ALPAFTTIOASBEGPAD-EGVFPAPANL--PARSEAAVQ-----PYIGISQVRMNSK 77
DB 431 AVIPEFTREFNLTINSFGHEHIVFPFGNEIYTPASNTLVVEVTLVPIETI----- 481
QY 78 EKDDGLTGYVLGITMWTIIAIGAGILGYSYKRGDL-----KEQHDQK-VCE 126
DB 482 -----LGIIG-----IVITAVFLKRSRQGTGFGPIKPELLTQDHDHTKQISV 526
QY 127 REMQR 131
DB 527 REAYR 531
RESULT 11
Q93J57 PRELIMINARY; PRT; 517 AA.
AC Q93J57;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative integral membrane protein.
GN SC00987 OR SCBAC19F3.14.
OS Streptomyces coelicolor.
OC Actinomycetes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.M., Challis G.L.,

RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleiser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleiser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S., Taylor K.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2)."
RL Nature 417:141-147(2002).
DR EMBL; AL596102; CAC44337.1; -
SQ SEQUENCE 517 AA; 54935 MW; 8C588592126A1F0 CRC64;
Query Match 9.1%; Score 83.5; DB 16; Length 517;
Best Local Similarity 28.7%; Pred. No. 16;
Matches 33; Conservative 15; Mismatches 40; Indels 27; Gaps 6;
QY 21 TSOALPAPFTT--EIQASBEGPAD-EGVFPAPANL-----ALPARSEAAVQVP 65
DB 323 TSGGLPAGLTVQVLEKARAGLDGFEIETPDDRSTWSVAQTDGLPVRMDQAAVDPH 382
QY 66 IG-ISQVRMNSK-----KKDGLTGYVLGITMWTIIAIGAGILGYSY 108
DB 383 TGEVTAPEWADHPVLAKLSTLGVRAHMGTLFGLANQIVLAIVLGLTGTIVLG 437
RESULT 12
Q9PA31 PRELIMINARY; PRT; 733 AA.
AC Q9PA31;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Primosomal protein N.
GN XF2689.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Brites M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carier H.,
RA Colaudo N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
RA Garner M., Goldman G.H., Goldman W.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hohnselt J.D., Junqueira M.L., Kemper E.L., Kitejima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Martino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nham A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmeri D.A., Paris A.,
RA Peixoto B.R., Pereira P.G., Rodrigues V., de Rosa A.U.M.,
RA Quaggio R.B., Roberto P.G., Rodriques V., de Rosa A.U.M.,
RA de Rosa V.E. Jr., de Sa R.G., Sanceli R.V., Sawaaki H.B.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
RA Vallada H., Van Sluys M.A., Verjovsky-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa."
RL Nature 406:151-159(2000).
CC -1- SIMILARITY: TO HELICASE C-TERMINAL DOMAIN.

DR EMBL; AE004075; AAF85486.1; -
 DR InterPro: IPR001650; Helicase_C.
 DR InterPro: IPR005259; P1A.
 DR Pfam; PF00271; helicase_C.1.
 DR SMART; SM00490; HELIC_C.1.
 DR TIGRfam; TIGR00595; p1A.1.
 KW ATP-binding; Helicase; Complete proteome.
 SQ SEQUENCE 733 AA; 81121 MW; E1341AFDDE759078 CRC64;

Query Match 9.1%; Score 83.5; DB 16; Length 733;
 Best Local Similarity 22.9%; Pred. No. 25;
 Matches 35; Conservative 17; Mismatches 52; Indels 49; Gaps 3;

QY 9 KRPEEDLRCEPTTSQALPAFTTEIOEASE-----GPGADEYQVAPAN 51
 DB 181 KKPPTVSLIIQTSTNTVPVANTEQKAVETLNANVGKTYLLDGVTSKTEVYLOAIAT 240
 QY 52 ALPARSEAAAVQPVIGISQVRNMSKEKKDL-----GTLGYV 88
 DB 241 CLAAKQALVLPVEIGLTPQLTNFPHALGI PVHALSKLDNERARVMAAARCGEARLV 300
 QY 89 LGITMMVILIIAIGAGIILGYSYKRGDKLKEQHD 121
 DB 301 LGTRSAVFPLPHAGLI-----VDEEHD 324

RESULT 13

Q95Q77 PRELIMINARY; PRT; 1553 AA.

AC Q95Q77; 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical 173.2 kDa protein.
 GN C32E12.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=9069613; PubMed=9851916;
 RA Waterston R.;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Wilcox L.;
 RT "The sequence of C. elegans cosmid C32E12.";
 RL submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RA Submitted (APR-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL; U80032; ALA16309.2; -
 KW Hypothetical protein.
 SQ SEQUENCE 1553 AA; 173196 MW; 02EAA041B6620CCE CRC64;

Query Match 9.1%; Score 83; DB 5; Length 1553;
 Best Local Similarity 18.1%; Pred. No. 70;
 Matches 36; Conservative 33; Mismatches 76; Indels 54; Gaps 5;

QY 3 EAGVPEKRCEDLRCE-----TTSQALPAFTTIOEASGSGADEYQV 46
 DB 354 KSSVETKTFETEKSEYQKATEPEKIDNFQSTPVTVPRGPNLVDTSSDEBEDVEI 411
 QY 47 FAPANALPARSEAAAVQPVIGISQVRNMSKEKKDLGTLGYLGTMMVILIIAIGAGIIL 106
 DB 412 IKPFIQ-PEKASEVLPVPIKTVSSIKINGIVKOV----- 446

QY 107 GYSYKRGDKLKEQHDQKVCEREMQRTILPLSAFTNPCEIVDEKTVV----HTSQTPV- 161
 DB 447 -----EKTEDKKKEVDKSKVATNIQDVPLPIPSFETSKVPENLISSAVKVLNHDGSIIPH 502
 QY 162 ----DROEGSTPLMGQAGT 176
 DB 503 EKSDDPQEFSEIKSSSESDT 521

RESULT 14

Q9LI28 PRELIMINARY; PRT; 774 AA.

AC Q9LI28; 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE ESTE D23839 (R0339).
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 RT clone: P0708G02.";
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AP001539; BAA92910.1; -
 DR InterPro: IPR004087; KH dom.
 DR InterPro: IPR004088; KH_type_1.
 DR Pfam; PF00013; KH-domain; 4.
 DR SMART; SM00322; KH; 4.
 DR PROSITE; PSS0084; KH TYPE 1; 4.
 SQ SEQUENCE 774 AA; 84066 MW; 5992CCFE0F74B56 CRC64;

Query Match 9.0%; Score 82; DB 10; Length 774;
 Best Local Similarity 23.0%; Pred. No. 37;
 Matches 29; Conservative 26; Mismatches 59; Indels 12; Gaps 4;

QY 41 ADEYQVAPANALP-----ARSEAAVQPVIGISQVRNMSKEKKDLGTLGYLGTMM 94
 DB 234 ADDSSVAKDINSRENAQLEKMPALVKAIVLPDRIFVNEHNGTGASGRNHSVR 293
 QY 95 VILIIAIGAGIILGYSYKRGDKLKEQHDQKVCEREMQRTILPLSAF-TNPTEIVDEK--T 151
 DB 294 LLVLDSSQVWLGG--KNGSVIKQWSTDSCEIRVSKDKLPLCALPRDELQILHOKIIT 350
 QY 152 VVVHTS 157
 DB 351 YSLHIS 356

RESULT 15

Q8W4W9 PRELIMINARY; PRT; 535 AA.

AC Q8W4W9; 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Transmembrane protein.
 GN pm3.
 OS Solanum tuberosum (Potato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4113;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. DESIREE; TISSUE=LEAF;
 RX MEDLINE=21576513; PubMed=11719809;

RA Rausch C., Daram P., Brunner S., Janas J., Laloi M., Leggewie G.,
 RA Amrhein N., Bucher M.,
 RT "A phosphate transporter expressed in arbuscule-containing cells in
 RT potato";
 RL Nature 414:462-466 (2001).
 DR EMBL; AJ318822; CAC87043.1; -.
 DR InterPro; IPR004738; Phos_permease.
 DR InterPro; IPR003662; sub_transporter.
 DR Pfam; PF00083; sugar_tr; 1.
 DR TIGRPFAM; TIGR00887; 2A0109; 1.
 SQ SEQUENCE 535 AA; 58565 MW; 94AF3CFC1839B1C1 CRC64;

Query Match 8.9%; Score 81.5; DB 10; Length 535;
 Best Local Similarity 25.8%; Pred. NO. 26;

Matches 32; Conservative 15; Mismatches 50; Indels 27; Gaps 5;

QY 35 ASRGPGADEVQVPAPANALPAR-----SEAAVQPVIGISQRYMNSKEKKDL 82
 DB 417 ANFGPNATTFV--PAEIFPARLRSTCHGISAAAGKAGIVGAFGLYAAQSTDPKVD- 473
 QY 83 GTIGYVIGITMTMTITTAIGAGIILGYSY-----KRGDKLKEQHDQVCEREMQRT---- 132
 DB 474 --AGYPTGIGVKNALIVLCVNFGLMFTLLVPSKSGKSLSEMSKENGEEMTKVENAQ 531
 QY 133 TLPL 136
 DB 532 TIPV 535

Search completed: April 7, 2003, 09:21:41
 Job time : 53.0062 secs

GenCore version 5.1.4 PJ.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: April 7, 2003, 09:16:24 ; Search time 33.4927 Seconds
(without alignments)
712.151 Million cell updates/sec

Title: US-10-057-951-2_COPY_85_263

Perfect score: 916
Sequence: 1 SGEAGVPRKPCEDLRCPET.....PVDQEGSTPLMGAGTPGA 179

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_101002: *
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT: *
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT: *
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT: *
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT: *
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT: *
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT: *
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT: *
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT: *
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT: *
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT: *
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT: *
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT: *
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: *
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT: *
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: *
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT: *
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: *
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT: *
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT: *
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT: *
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: *
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: *
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	916	100.0	263	20	AAV05219
2	916	100.0	263	22	AAW87769
3	916	100.0	263	22	AAE00300
4	913	99.7	263	21	AAAB4237
5	913	99.7	263	22	AAU086149
6	902	98.5	263	22	AAW93748
7	704.5	76.9	286	22	AAV05220
8	322	35.2	66	22	ABB37905
9	322	35.2	66	22	ABB23159
10	322	35.2	66	22	AAW58537

11	322	35.2	66	22	AAW1037	Human bone marrow
12	322	35.2	66	22	AAW18800	Peptide #5234 enco
13	322	35.2	66	22	AAW31314	Peptide #5351 enco
14	322	35.2	66	22	ABG40828	Human peptide enco
15	95	10.4	163	22	AAU58844	Propionibacterium
16	94	10.3	8805	22	ABR67112	Drosophila melanog
17	86.5	9.4	507	20	AAV34756	C. pneumoniae prot
18	80.5	8.8	769	21	AAV53896	A. Neisseria mening
19	80	8.7	1338	21	ABW07823	Human notch agonis
20	79.5	8.7	769	21	AAV53897	A. Neisseria mening
21	79	8.6	1212	19	AAW44299	Human serrate 2.
22	78.5	8.6	552	22	ABW71591	Drosophila melanog
23	77.5	8.5	452	22	AAW82370	S. epidermidis ope
24	77.5	8.5	463	23	ABP38128	Staphylococcus epi
25	77.5	8.5	595	23	AAU78088	Human CD30 protein
26	77	8.4	769	21	AAV53895	A. Neisseria mening
27	76.5	8.4	425	21	AAV92370	p75-NTR (neurotrop
28	76.5	8.4	425	23	AAE21671	Rat neurotrophic r
29	76.5	8.4	461	19	AAW75391	Human liver glyco
30	76.5	8.4	501	19	AAW75391	Human liver glyco
31	76.5	8.4	556	22	ABG00252	Novel human diagno
32	76.5	8.4	556	22	ABG00603	Novel human diagno
33	76.5	8.4	556	22	ABG03054	Novel human diagno
34	76.5	8.4	1155	22	AAW40061	Human polypeptide
35	75.5	8.2	219	23	ABP40806	Staphylococcus epi
36	75.5	8.2	276	22	ABW31724	Peptide #4375 enco
37	75.5	8.2	276	22	ABW31724	Peptide #4375 enco
38	75.5	8.2	276	22	ABW31724	Peptide #4375 enco
39	75.5	8.2	276	22	AAW70088	Human bone marrow
40	75.5	8.2	276	22	AAW17915	Peptide #4349 enco
41	75.5	8.2	276	22	AAW30426	Peptide #4463 enco
42	75.5	8.2	276	22	AAW05567	Peptide #4249 enco
43	75.5	8.2	276	22	ABG39722	Human peptide enco
44	75	8.2	665	22	ABG15591	Novel human diagno
45	75	8.2	595	14	AAW35478	Lymphocyte activat
					AAW24017	Human soluble CD30

ALIGNMENTS

RESULT 1	AAV05219	standard; Protein; 263 AA.
ID	AAV05219;	
AC	AAV05219;	
XX	17-JUN-1999 (first entry)	
DT	17-JUN-1999 (first entry)	
XX	Kirngliel protein sequence.	
DE	Kirngliel protein sequence.	
XX	Kirngliel cancer; inflammation; autoimmunity; rheumatoid arthritis; AIDS;	
KW	CNS inflammation; cerebellar degeneration; Alzheimer's disease; asthma;	
KW	Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;	
KW	neurological abnormality; ischemia reperfusion injury; ischaemic injury;	
KW	cardiovascular disease; kidney disease; liver disease; aplastic anaemia;	
KW	myocardial infarction; hypotension; hypertension; allergy; infection;	
KW	myelodysplastic syndrome; haematologic abnormality; diagnosis; therapy;	
KW	male pattern baldness.	
XX	Homo sapiens.	
OS	Homo sapiens.	
XX	WO9911788-A1.	
PN	WO9911788-A1.	
XX	11-MAR-1999.	
PD	11-MAR-1999.	
XX	02-SEP-1998; 98WO-US18270.	
XX	02-SEP-1998; 98US-014489.	
PR	01-SEP-1998; 98US-014489.	
XX	02-SEP-1997; 97US-0056032.	
XX	(SMIK) SMITHKLINE BEECHAM CORP.	
PA	(SMIK) SMITHKLINE BEECHAM CORP.	
XX	Albone EF, Kikly KK;	
PI	Albone EF, Kikly KK;	

```
XX WP1: 1999-214707/18.
DR N-PSDB; AAX28354.
XX
XX New kringlet polypeptides and polynucleotides
XX
XX Claim 1; Page 31-32; 42pp; English.
XX
XX This sequence is a Kringlet polypeptide of the invention.
XX The kringlet polypeptides (I) are used to screen for agonists and
XX antagonists. Agonists are used to treat subjects in need of enhanced
XX activity or expression of (I). Antagonists are used to treat subjects
XX having need to inhibit the activity or expression of (I). The methods can
XX be used to treat conditions such as cancer, inflammation, autoimmunity,
XX allergy, asthma, rheumatoid arthritis, CNS inflammation, cerebellar
XX degeneration, Alzheimer's disease, Parkinson's disease, multiple
XX sclerosis, amyotrophic lateral sclerosis, head injury damage and other
XX neurological abnormalities, liver disease, ischaemic injury, myocardial
XX infarction, hypotension, hypertension, AIDS, myelodysplastic syndromes
XX and other haematologic abnormalities, aplastic anaemia, male pattern
XX baldness, and bacterial, fungal, protozoan and viral infections. The
XX kringlet polypeptides may also be used to generate antibodies.
XX Determining the presence or absence of mutations in, and analysing for
XX the presence or absence of expression of, kringlet polynucleotides can be
XX used to diagnose a disease or susceptibility to a disease related to
XX expression or activity of kringlet proteins. The polynucleotides may also
XX be used for chromosome identification, and mapping.
XX
XX Sequence 263 AA;
SQ
Query Match 100.0%; Score 916; DB 20; Length 263;
Best Local Similarity 100.0%; Pred. No. 2,3e-89;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIOEASSEGGADEVQVAPANALPARSEAA 60
DB 85 SGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIOEASSEGGADEVQVAPANALPARSEAA 144
QY 61 AVQPIVIGISQVRNMSKSKDGLTGIVLGIITMMVITIIAGAGIILGYSTKRGDKLEQH 120
DB 145 AVQPIVIGISQVRNMSKSKDGLTGIVLGIITMMVITIIAGAGIILGYSTKRGDKLEQH 204
QY 121 DQKVCEREMQRTITPLSAFTNPTECEIVDEKTVVHTSQTVPDPOEGSTPIPMGAGTPGA 179
DB 205 DQKVCEREMQRTITPLSAFTNPTECEIVDEKTVVHTSQTVPDPOEGSTPIPMGAGTPGA 263
RESULT 2
AAW87769
ID AAW87769 standard; Protein; 263 AA.
XX
XX AAW87769;
AC
XX
XX 29-MAR-1999 (first entry)
DE
XX
XX Human tissue plasminogen activator-like protease t-PALP.
XX
XX Tissue plasminogen activator-like protease; t-PALP; human;
XX circulatory system-related disorder; blood clotting; stroke;
XX thrombosis; peripheral arterial occlusion; pulmonary embolism;
XX myocardiothrombosis; diagnosis; therapy.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Peptide 1..21
FT /label= sig_peptide
FT 22..263
FT /label= Mat_protein
FT Domain 25..84
FT /note= "kringle domain"
FT Domain 85..263
```

```
FT Peptide /note= "protease domain"
FT 22..31
FT /note= "epitope-bearing region"
FT 35..44
FT /note= "epitope-bearing region"
FT 71..81
FT /note= "epitope-bearing region"
FT 91..107
FT /note= "epitope-bearing region"
FT 119..128
FT /note= "epitope-bearing region"
FT 138..147
FT /note= "epitope-bearing region"
FT 155..167
FT /note= "epitope-bearing region"
FT 193..203
FT /note= "epitope-bearing region"
FT 206..215
FT /note= "epitope-bearing region"
FT 227..237
FT /note= "epitope-bearing region"
FT 243..252
FT /note= "epitope-bearing region"
XX
XX W09854199-A1.
XX
XX 03-DEC-1998.
XX
XX 27-MAY-1998; 98MO-US10728.
XX
XX 28-MAY-1997; 97US-0048000.
XX
XX (HUMA-) HUMAN GENOMEG SCI INC.
XX
XX Ebner R, Moore PA, Ruben SM;
XX
XX WPI: 1999-070207/06.
XX
XX N-PSDB; AAV99636.
XX
XX New tissue plasminogen activator-like protease - useful in the
XX diagnosis and treatment of circulatory system-related disorders
XX
XX Claim 1; Page 56-57; 76pp; English.
XX
XX This is the amino acid sequence of tissue plasminogen activator-like
XX protease (t-PALP), a novel member of the serine protease family
XX that shares sequence homology to human tissue plasminogen activator
XX (see AAW87770). The t-PALP sequence was deduced from a cDNA clone
XX (see AAV99636) derived from activated monocytes. The 2.5 kb t-PALP
XX message has also been detected in heart, brain, lung, placenta,
XX liver, skeletal muscle, kidney, pancreas, spleen, thymus, prostate,
XX testis, ovary, small intestine, colon and peripheral blood
XX leukocytes. Isolated nucleic acids encoding amino acids -21 to
XX 242, -20 to 242, 1-242, 4-63 (kringle domain) and 64-242 (protease
XX domain) of t-PALP, or encoding epitope-bearing portions of t-PALP,
XX are also claimed, as are recombinant vectors, host cells, and
XX methods for producing t-PALP polypeptides. t-PALP may be used to
XX detect and treat disorders related to the circulatory system, and
XX to identify agonists and antagonists of t-PALP activity. The
XX homology between t-PALP and tPA indicates that t-PALP may be
XX involved in the regulation of normal and abnormal clotting
XX in e.g. stroke, deep-vein thrombosis, peripheral arterial
XX occlusion, pulmonary embolism and myocardiothrombosis.
XX
XX Sequence 263 AA;
SQ
Query Match 100.0%; Score 916; DB 20; Length 263;
Best Local Similarity 100.0%; Pred. No. 2,3e-89;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIOEASSEGGADEVQVAPANALPARSEAA 60
DB 85 SGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIOEASSEGGADEVQVAPANALPARSEAA 144
```


PR 05-JAN-2000; 2000WO-US00219.
XX (GETH) GENENTECH INC.
XX
XX Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
PI Marsters SA, Pan J, Picti RM, Roy MA, Smith V, Stone DM;
PI Matarabe CK, Wood WI.
XX WPI; 2002-205567/26.
DR N-PSDB; ABR40275.
XX
XX Thirty five nucleic acids encoding PRO polypeptides, useful for
PT treating benign or malignant tumours, leukaemias and lymphoid
PT malignancies, inflammatory, angiogenic and immunologic disorders -
XX
XX Claim 61; Fig 44; 302pp; English.
XX
XX The present invention relates to the isolation of novel human PRO
CC polypeptides and the polynucleotide sequences encoding them. The
CC PRO polypeptides, agonists, antagonists or anti-PRO antibodies are
CC useful for treating benign or malignant tumours (e.g. renal, kidney,
CC bladder, breast, etc), leukaemias and lymphoid malignancies, other
CC disorders such as neuronal, glial, astrocytal, hypothalamic, glandular,
CC macrophagal, stromal and blastocoeleic disorders, inflammatory, immune
CC and angiogenic disorders. The polynucleotide sequences are also
CC useful in gene therapy. AAU86128-AAU86162 represent the human PRO
CC polypeptides of the invention.
XX
XX Sequence 263 AA;
SQ
Query Match 99.7%; Score 913; DB 23; Length 263;
Best Local Similarity 99.4%; Pred. No. 4.8e-89;
Matches 178; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SGEAGVPEKRPCEDLRCPEETSQALPAFTTEIOEASGPGADEVQVAPANALPARSEAA 60
DB 85 SGEAGVPEKRPCEDLRCPEETSQALPAFTTEIOEASGPGADEVQVAPANALPARSEAA 144
QY 61 AVQPVIGISQVRMNSKSKKDLGTLGVVLGITMMVIIAIGAGIILGYSKRGDKLEQH 120
DB 145 AVQPVIGISQVRMNSKSKKDLGTLGVVLGITMMVIIAIGAGIILGYSKRGDKLEQH 204
QY 121 DQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTSQTPVDPEGSTPLMGAGTPGA 179
DB 205 DQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTSQTPVDPEGSTPLMGAGTPGA 263
RESULT 6
AAU93748
ID AAM93748 standard; Protein; 263 AA.
XX
XX AAM93748;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human polypeptide, SEQ ID NO: 3727.
XX
XX Human; full length cDNA; cDNA synthesis; oligo-capping.
XX
XX Homo sapiens.
XX
XX EP1130094-A2.
XX
XX 05-SEP-2001.
XX
XX 07-JUL-2000; 2000EP-0114089.
XX
XX 08-JUL-1999; 99JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0183765.
XX
XX (HELI-) HELIX RES INST.
XX

PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI; 2001-524255/58.
DR N-PSDB; AAK94700.
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
XX Claim 8; SEQ ID NO 3727; 1380pp + sequence listing; English.
XX
XX The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a polypeptide
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
XX Sequence 263 AA;
SQ
Query Match 98.5%; Score 902; DB 22; Length 263;
Best Local Similarity 98.9%; Pred. No. 7.2e-88;
Matches 177; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 SGEAGVPEKRPCEDLRCPEETSQALPAFTTEIOEASGPGADEVQVAPANALPARSEAA 60
DB 85 SGEAGVPEKRPCEDLRCPEETSQALPAFTTEIOEASGPGADEVQVAPANALPARSEAA 144
QY 61 AVQPVIGISQVRMNSKSKKDLGTLGVVLGITMMVIIAIGAGIILGYSKRGDKLEQH 120
DB 145 AVQPVIGISQVRMNSKSKKDLGTLGVVLGITMMVIIAIGAGIILGYSKRGDKLEQH 204
QY 121 DQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTSQTPVDPEGSTPLMGAGTPGA 179
DB 205 DQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTSQTPVDPEGSTPLMGAGTPGA 263
RESULT 7
AAU05220
ID AAY05220 standard; Protein; 286 AA.
XX
XX AAY05220;
XX
XX 17-JUN-1999 (first entry)
XX
XX Kringel protein sequence.
XX
XX Kringel; cancer; inflammation; autoimmunity; rheumatoid arthritis; AIDS;
XX CNS inflammation; cerebellar degeneration; Alzheimer's disease; asthma;
XX Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
XX neurological abnormality; ischaemia reperfusion injury; ischaemic injury;
XX cardiovascular disease; kidney disease; liver disease; aplastic anaemia;
XX myocardial infarction; hypotension; hypertension; allergy; infection;
XX myelodysplastic syndrome; haematologic abnormality; diagnosis; therapy;
XX male pattern baldness.
XX
XX Homo sapiens.
XX
XX WO9911788-A1.
XX
XX 11-MAR-1999.
XX
XX 02-SEP-1998; 98WO-US18270.
XX
XX 01-SEP-1998; 98US-0144889.
PR 02-SEP-1997; 97US-0056032.
XX
XX

PA (SMIK) SMITHKLINE BEECHAM CORP.
XX Albione EF, Kirkly KK;
XX WPI; 1999-214707/18.
DR N-PSDB; AAX28355.
XX
PT New kringlet polypeptides and polynucleotides
XX
PS Claim 14; Page 33; 42pp; English.
XX
CC This sequence is a kringlet polypeptide of the invention.
CC The kringlet polypeptides (I) are used to screen for agonists and
CC antagonists. Agonists are used to treat subjects in need of enhanced
CC activity or expression of (I). Antagonists are used to treat subjects
CC having need to inhibit the activity or expression of (I). The methods can
CC be used to treat conditions such as cancer, inflammation, cerebellar
CC degeneration, Alzheimer's disease, Parkinson's disease, multiple
CC sclerosis, amyotrophic lateral sclerosis, head injury damage and other
CC neurological abnormalities, ischaemia reperfusion injury, cardiovascular
CC disease, kidney disease, liver disease, ischaemic injury, myocardial
CC infarction, hypertension, AIDS, myelodysplastic syndromes
CC and other haematologic abnormalities, aplastic anaemia, male pattern
CC baldness, and bacterial, fungal, protozoan and viral infections. The
CC kringlet polypeptides may also be used to generate antibodies.
CC Determining the presence or absence of mutations in, and analysing for
CC the presence or absence of expression of, kringlet polynucleotides can be
CC used to diagnose a disease or susceptibility to a disease related to
CC expression or activity of kringlet proteins. The polynucleotides may also
CC be used for chromosome identification, and mapping.
XX
SQ Sequence 286 AA;

Query Match 76.9%; Score 704.5; DB 20; Length 286;
Best Local Similarity 82.4%; Pred. No. 9.9e-67;
Matches 145; Conservative 5; Mismatches 25; Indels 1; Gaps 1;

QY 1 SGEAGVPEKRPCEBDRCPETTSQALPAF-TTEIOEASGEGADVOVFAPANALPARSEA 59
DB 85 SGEAGVPEKRPCEBDRCPETTSQALPAFHDRNSKRLEGGADEVGFAPANALPARSEA 144

QY 60 AAVQPVIGISQRYVMSKEKKDLGTLGYVLITMMVLIIGAGIILGYSKRGKDLKEQ 119
DB 145 ALLOPVIGISQRYVDEQGEKRGNSGLRAHGHDIILIGAGIILGYSKRGKDLKEQ 204

QY 120 HDQVCEREMQRTLPISAFNPTCEIIVDEKTVVHTSQTPVDPQEGSTPLMGOAG 175
DB 205 HDQVCEREMQRTLPISAFNPTCEIIVDEKTVVHTSQTPVDPQEGSTPLMGOAG 260

RESULT 8
ABB37905
ID ABB37905 standard; Peptide; 66 AA.
XX
AC ABB37905;
XX
DT 04-FEB-2002 (first entry)
XX
DE Peptide #5411 encoded by human foetal liver single exon probe.
XX
KM Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001MO-US00669.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-02344687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human fetal liver -
XX
PS Claim 27; SEQ ID NO 30540; 639pp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC foetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 66 AA;

Query Match 35.2%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 8.9e-27;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 TTSQLPAFTTEIOBASGPGADVOVFAPANALPARSEAAAVPVIGISQRYVMSKEK 79
DB 1 TTSQLPAFTTEIOBASGPGADVOVFAPANALPARSEAAAVPVIGISQRYVMSKEK 60

QY 80 KDLGTL 85
DB 61 KDLGTL 66

RESULT 9
ABB23159
ID ABB23159 standard; Protein; 66 AA.
XX
AC ABB23159;
XX
DT 23-JAN-2002 (first entry)
XX
DE Protein #5158 encoded by probe for measuring heart cell gene expression.
XX
XX
KM Human; gene expression; heart; microarray; vascular system;
KM cardiovascular disease; hypertension; cardiac arrhythmia;
KM congenital heart disease.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001MO-US00666.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-02344687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

```
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX XX
DR WPI; 2001-48899/53.
XX PT Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX PS Claim 15; SEQ ID No 24929; 530pp; English.
XX CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21533-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarray.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosticating diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 66 AA;
XX
Query Match 35.2%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 8.9e-27;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 20 TTSQLPAPFTTEIOEASBGPADDEVQVAPANALPARSEAAAVQVIGISQVRMNSKEK 79
DB 1 TTSQLPAPFTTEIOEASBGPADDEVQVAPANALPARSEAAAVQVIGISQVRMNSKEK 60
XX
QY 80 KDLGTL 85
DB 61 KDLGTL 66
XX
RESULT 10
AAM58537
ID AAM58537 standard; Protein; 66 AA.
XX
AC AAM58537;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 30642.
XX
KM Human; brain expressed exon; gene expression analysis; probe;
KM microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KM epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN MO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX
```

```
XX PT Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX PS Example 4; SEQ ID NO: 30642; 650pp + Sequence Listing; English.
XX CC The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is a protein encoded by one of
XX the probes of the invention.
XX SQ Sequence 66 AA;
XX
Query Match 35.2%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 8.9e-27;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 20 TTSQLPAPFTTEIOEASBGPADDEVQVAPANALPARSEAAAVQVIGISQVRMNSKEK 79
DB 1 TTSQLPAPFTTEIOEASBGPADDEVQVAPANALPARSEAAAVQVIGISQVRMNSKEK 60
XX
QY 80 KDLGTL 85
DB 61 KDLGTL 66
XX
RESULT 11
AAM71037
ID AAM71037 standard; Protein; 66 AA.
XX
AC AAM71037;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 31343.
XX
KM Human; bone marrow expressed exon; gene expression analysis; probe;
KM microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN MO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 31343; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
```

CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.

XX
SQ Sequence 66 AA;

Query Match 35.2%; Score 322; DB 22; Length 66;
Best Local Similarity 100.0%; Pred. No. 8.9e-27;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 TTSQLPAPFTTEIOEASGPGADEVQVFAPANALPARSEAAAVQPVIGISQVRVMSKKEK 79
DB 1 TTSQLPAPFTTEIOEASGPGADEVQVFAPANALPARSEAAAVQPVIGISQVRVMSKKEK 60

QY 80 KDLGTL 85
DB 61 KDLGTL 66

RESULT 12

AAM1800
ID AAM1800 standard; Protein; 66 AA.

XX AAM1800;

DT 12-OCT-2001 (first entry)

XX Peptide #5234 encoded by probe for measuring cervical gene expression.

XX KW Probe; human; microarray; gene expression; cervical epithelial cell;
XX cervical cancer.

XX OS Homo sapiens.

PN W0200157278-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00670.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488901/53.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human cervical epithelial cells -

XX Claim 27; SEQ ID No 23626; 487bp; English.

XX The present invention relates to human single exon nucleic acid probes

XX (SENPs: see AAI10068-AI12459). The present sequence is a peptide encoded

XX by one such probe. The SENPs are derived from human HeLa cells. The SENPs

XX can be used to produce a single exon microarray, which can be used for

XX measuring human gene expression in a sample derived from human cervical

XX epithelial cells. By measuring gene expression, the probes are therefore

XX useful in grading and/or staging of diseases of the cervix, notably

XX cervical cancer.

XX Note: The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at fcp.wipo.int/pub/published_pct_sequences.

XX Sequence 66 AA;

XX Query Match 35.2%; Score 322; DB 22; Length 66;

Best Local Similarity 100.0%; Pred. No. 8.9e-27;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 TTSQLPAPFTTEIOEASGPGADEVQVFAPANALPARSEAAAVQPVIGISQVRVMSKKEK 79
DB 1 TTSQLPAPFTTEIOEASGPGADEVQVFAPANALPARSEAAAVQPVIGISQVRVMSKKEK 60

QY 80 KDLGTL 85
DB 61 KDLGTL 66

RESULT 13

AAM31314
ID AAM31314 standard; Protein; 66 AA.

XX AAM31314;

DT 17-OCT-2001 (first entry)

XX Peptide #5351 encoded by probe for measuring placental gene expression.

XX KW Probe; microarray; human; placenta; antenatal diagnosis;

XX genetic disorder.

XX OS Homo sapiens.

PN W0200157272-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00663.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488897/53.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human placenta -

XX Claim 27; SEQ ID No 31583; 654bp; English.

XX The present invention relates to single exon nucleic acid probes (SENPs:

XX see AAI31315-AI157546). The present sequence is a peptide encoded by one

XX such probe. The probes are useful for producing a microarray for

XX predicting, measuring and displaying gene expression in samples derived

XX from human placenta. The probes are useful for antenatal diagnosis of

XX human genetic disorders.

XX Sequence 66 AA;

XX Query Match 35.2%; Score 322; DB 22; Length 66;

XX Best Local Similarity 100.0%; Pred. No. 8.9e-27;
XX Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 TTSQLPAPFTTEIOEASGPGADEVQVFAPANALPARSEAAAVQPVIGISQVRVMSKKEK 79
DB 1 TTSQLPAPFTTEIOEASGPGADEVQVFAPANALPARSEAAAVQPVIGISQVRVMSKKEK 60

QY 80 KDLGTL 85
DB 61 KDLGTL 66

RESULT 14
 ABG40828
 ID ABG40828 standard; Peptide; 66 AA.
 AC
 XX ABG40828;
 XX
 DT 19-AUG-2002 (first entry)
 DE Human peptide encoded by genome-derived single exon probe SEQ ID 30493.
 XX
 XX Human; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagazer syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX
 OS Homo sapiens.
 XX
 XX MO200186003-A2.
 PN
 PD 15-NOV-2001.
 XX
 XX 30-JAN-2001; 2001WO-US00665.
 PF
 XX
 XX 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI; 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 PS
 PS Claim 27; SEQ ID No 30493; 634pp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung; comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryotic lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probe/open reading frames (ORF). The probes are used for gene

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:19:28 ; Search time 12.6528 Seconds
(without alignments)
864.895 Million cell updates/sec

Title: US-10-057-951-2_COPY_85_263

Sequence: 1 SGBAGVPEKRPCEDLRCPERT.....PVDPQSGTFLMGQAGTPGA 179

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues
Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/1/pubpaa/PC7_NEW_PUB pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/1/pubpaa/FC105_PUBCOMB pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Match Length	ID	Description
1	916	100.0	263 9 US-10-102-704-2	Sequence 2, Appl1
2	916	100.0	263 9 US-10-057-951-2	Sequence 2, Appl1
3	916	100.0	263 10 US-09-084-491A-2	Sequence 2, Appl1
4	322	35.2	66 10 US-09-864-761-38457	Sequence 38457, A
5	80	8.7	1212 9 US-10-219-248-3	Sequence 3, Appl1
6	80	8.7	1212 9 US-10-219-247-3	Sequence 3, Appl1
7	80	8.7	1212 10 US-09-855-722-3	Sequence 3, Appl1
8	80	8.7	1238 9 US-10-219-248-5	Sequence 5, Appl1
9	80	8.7	1238 10 US-10-219-247-5	Sequence 5, Appl1
10	80	8.7	1238 9 US-09-855-722-5	Sequence 5, Appl1
11	80	8.7	1238 10 US-09-944-849-4	Sequence 6, Appl1
12	77.5	8.5	1709 9 US-09-921-667-6	Sequence 6, Appl1
13	77.5	8.5	1709 9 US-09-870-759-51	Sequence 5, Appl1
14	76.5	8.4	425 10 US-09-746-537-14	Sequence 14, Appl1
15	76.5	8.4	425 10 US-09-821-831-2	Sequence 2, Appl1
16	75.5	8.2	276 10 US-09-864-761-46690	Sequence 46690, A
17	75	8.2	595 10 US-09-826-212-9	Sequence 9, Appl1
18	75	8.2	595 10 US-09-935-727-11	Sequence 11, Appl1
19	74.5	8.1	403 10 US-09-996-194-6	Sequence 6, Appl1

20	74.5	8.1	473 10 US-09-764-864-802	Sequence 802, App
21	74.5	8.1	617 9 US-09-860-670-82	Sequence 82, Appl1
22	74.5	8.1	617 10 US-09-764-864-1263	Sequence 1263, Ap
23	73.5	8.0	1179 10 US-09-821-883-29	Sequence 29, Appl1
24	73	8.0	348 10 US-09-731-872-247	Sequence 247, App
25	73	8.0	451 10 US-09-935-390A-26	Sequence 26, Appl1
26	73	8.0	461 9 US-09-992-598-301	Sequence 301, App
27	73	8.0	461 9 US-09-989-293A-301	Sequence 301, App
28	73	8.0	461 9 US-09-989-735-301	Sequence 301, App
29	73	8.0	461 9 US-09-990-444-301	Sequence 301, App
30	73	8.0	461 9 US-09-989-730-301	Sequence 301, App
31	73	8.0	461 9 US-09-990-436-301	Sequence 301, App
32	73	8.0	461 9 US-09-991-181-301	Sequence 301, App
33	73	8.0	461 9 US-09-993-687-301	Sequence 301, App
34	73	8.0	461 9 US-09-989-734-301	Sequence 301, App
35	73	8.0	461 9 US-10-028-072-454	Sequence 454, App
36	73	8.0	461 9 US-09-997-653-301	Sequence 301, App
37	73	8.0	461 9 US-09-993-667-301	Sequence 301, App
38	73	8.0	461 9 US-10-121-049-454	Sequence 454, App
39	73	8.0	461 9 US-10-123-904-454	Sequence 454, App
40	73	8.0	461 9 US-10-140-470-454	Sequence 454, App
41	73	8.0	461 9 US-09-990-438-301	Sequence 301, App
42	73	8.0	461 9 US-09-990-562-301	Sequence 301, App
43	73	8.0	461 9 US-09-997-428-301	Sequence 301, App
44	73	8.0	461 9 US-09-997-666-301	Sequence 301, App
45	73	8.0	461 9 US-10-175-746-454	Sequence 454, App

ALIGNMENTS

```
RESULT 1
US-10-102-704-2
; Sequence 2, Application US/10102704
; Parent No. US200201647688A1
; GENERAL INFORMATION:
; APPLICANT: Moore et al.
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protein
; FILE REFERENCE: PF378C1
; CURRENT APPLICATION NUMBER: US/10/102,704
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/084,491
; PRIOR FILING DATE: 1998-05-27
; PRIOR APPLICATION NUMBER: 60/048,000
; PRIOR FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-102-704-2

Query Match      100.0%; Score 916; DB 9; Length 263;
Best Local Similarity 100.0%; Pred. No. 4.9e-87;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGBAGVPEKRPCEDLRCPERTTSQALPAFTTEIOEASEGGADEVQVFAFANALPARSEAA 60
DB 85 SGBAGVPEKRPCEDLRCPERTTSQALPAFTTEIOEASEGGADEVQVFAFANALPARSEAA 144

QY 61 AVQPVIGISQVRNNSKKKDLGTLGVTWVIAIAGGIIIGSYRKGDLEKQH 120
DB 145 AVQPVIGISQVRNNSKKKDLGTLGVTWVIAIAGGIIIGSYRKGDLEKQH 204

QY 121 DQKVCERMRITLPLSAFTNPCEIVDEKTVVHTSOPVDPQSGTFLMGQAGTPGA 179
DB 205 DQKVCERMRITLPLSAFTNPCEIVDEKTVVHTSOPVDPQSGTFLMGQAGTPGA 263

RESULT 2
US-10-057-951-2
; Sequence 2, Application US/10057951
```

Parent No. US20020177213A1
GENERAL INFORMATION:
APPLICANT: Moore et al.
TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
FILE REFERENCE: PF378P1
CURRENT APPLICATION NUMBER: US/10/057,951
CURRENT FILING DATE: 2002-01-29
PRIOR APPLICATION NUMBER: US 09/411,977
PRIOR FILING DATE: 1999-10-04
PRIOR APPLICATION NUMBER: US 09/084,491
PRIOR FILING DATE: 1998-05-27
PRIOR APPLICATION NUMBER: US 60/048,000
PRIOR FILING DATE: 1997-05-28
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 2
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
US-10-057-951-2

Query Match 100.0%; Score 916; DB 9; Length 263;
Best Local Similarity 100.0%; Pred. No. 4,9e-87;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGEAGVPEKRCPCEDLRCPETTSQALPAFTTEIOEASGEGADEVGFAPANALPARSEAA 60
DB 85 SGEAGVPEKRCPCEDLRCPETTSQALPAFTTEIOEASGEGADEVGFAPANALPARSEAA 144
QY 61 AVOPVIGISQVRNNSKEKDLGLGVIGITMMVILIAIGAILLGSYKRGDLKEQH 120
DB 145 AVOPVIGISQVRNNSKEKDLGLGVIGITMMVILIAIGAILLGSYKRGDLKEQH 204
QY 121 DQVCEREMQRTILPLSAFTNPCEIVDEKTVVHTSQTPVDPQGSTPLMGAGTPGA 179
DB 205 DQVCEREMQRTILPLSAFTNPCEIVDEKTVVHTSQTPVDPQGSTPLMGAGTPGA 263

RESULT 3

US-09-084-491A-2
Sequence 2, Application US/09084491A
Patent No. US2002061576A1
GENERAL INFORMATION:
APPLICANT: MOORE, PAUL A.
APPLICANT: RUBEN, STEVEN M.
APPLICANT: EBNER, REINHARD
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/084,491A
FILING DATE: 27-MAY-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BROOKS, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF378
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

LENGTH: 263 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-084-491A-2

Query Match 100.0%; Score 916; DB 10; Length 263;
Best Local Similarity 100.0%; Pred. No. 4,9e-87;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGEAGVPEKRCPCEDLRCPETTSQALPAFTTEIOEASGEGADEVGFAPANALPARSEAA 60
DB 85 SGEAGVPEKRCPCEDLRCPETTSQALPAFTTEIOEASGEGADEVGFAPANALPARSEAA 144
QY 61 AVOPVIGISQVRNNSKEKDLGLGVIGITMMVILIAIGAILLGSYKRGDLKEQH 120
DB 145 AVOPVIGISQVRNNSKEKDLGLGVIGITMMVILIAIGAILLGSYKRGDLKEQH 204
QY 121 DQVCEREMQRTILPLSAFTNPCEIVDEKTVVHTSQTPVDPQGSTPLMGAGTPGA 179
DB 205 DQVCEREMQRTILPLSAFTNPCEIVDEKTVVHTSQTPVDPQGSTPLMGAGTPGA 263

RESULT 4

US-09-864-761-38457
Sequence 38457, Application US/09864761
Patent No. US2002048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Acemica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29

US-09-855-722-3

Query Match 8.7%; Score 80; DB 10; Length 1212;
Best Local Similarity 22.4%; Pred. No. 13;
Matches 55; Conservative 20; Mismatches 77; Indels 94; Gaps 11;

QY 2 GEAGV--PEKRPC-----EDLRC-----PETTS-----QALPAFTTEIQ 33
DB 921 GECGAEEPPSTPCLPBSGHLDDNNCARLTLHFNRDHVPQGTTVGALCSGIRSLPATRAVAR 980
QY 34 E-----ASEGPGADEVOV--FAPANALPARS--EAAAVQPVIGISQVRNMSKEKD 81
DB 981 DRLVLCDRASSGASAVEAVSFPARDLPDSSLICGAHAIVAATOR----- 1030
QY 82 LGTLGYVLGITMVTIITAIGAGIILG-----YSYRGKDLKE 118
DB 1031 -GNSSLLAVTEVKETVVTGSSSTGLVPLCGAFSVLMACVLCVWTRKRRK--- 1085
QY 119 QHDQKVCEREMQRT-----LPLSAFTNPTCEIVDEKTVVHTSQTVPDQEGSTP 169
DB 1086 -----ERERSRLPRESANNQWAPLNPRIERNRPGHKGDLVYQCKNFTPPRRADEA 1138
QY 170 LMGQAG 175
DB 1139 LGGPAG 1144

RESULT 8

US-10-219-248-5
; Sequence 5, Application US/10219248
; Publication No. US200300223688A1
; GENERAL INFORMATION:
; APPLICANT: Sakano, Sei-ji
; APPLICANT: Itoh, Akira
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-8576
; CURRENT APPLICATION NUMBER: US/10/219,248
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US/09/214,278
; PRIOR FILING DATE: 1999-01-26
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1238
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-219-248-5

Query Match 8.7%; Score 80; DB 9; Length 1238;
Best Local Similarity 22.4%; Pred. No. 13;
Matches 55; Conservative 20; Mismatches 77; Indels 94; Gaps 11;

QY 2 GEAGV--PEKRPC-----EDLRC-----PETTS-----QALPAFTTEIQ 33
DB 947 GECGAEEPPSTPCLPBSGHLDDNNCARLTLHFNRDHVPQGTTVGALCSGIRSLPATRAVAR 1006
QY 34 E-----ASEGPGADEVOV--FAPANALPARS--EAAAVQPVIGISQVRNMSKEKD 81
DB 1007 DRLVLCDRASSGASAVEAVSFPARDLPDSSLICGAHAIVAATOR----- 1056
QY 82 LGTLGYVLGITMVTIITAIGAGIILG-----YSYRGKDLKE 118
DB 1057 -GNSSLLAVTEVKETVVTGSSSTGLVPLCGAFSVLMACVLCVWTRKRRK--- 1111
QY 119 QHDQKVCEREMQRT-----LPLSAFTNPTCEIVDEKTVVHTSQTVPDQEGSTP 169
DB 1112 -----ERERSRLPRESANNQWAPLNPRIERNRPGHKGDLVYQCKNFTPPRRADEA 1164
QY 170 LMGQAG 175
DB 1165 LGGPAG 1170

RESULT 9

US-10-219-247-5
; Sequence 5, Application US/10219247
; Publication No. US20030032781A1
; GENERAL INFORMATION:
; APPLICANT: Sakano, Sei-ji
; APPLICANT: Itoh, Akira
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-8576
; CURRENT APPLICATION NUMBER: US/10/219,247
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US/09/855,722
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/214,278
; PRIOR FILING DATE: 1999-01-26
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1238
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-219-247-5

Query Match 8.7%; Score 80; DB 9; Length 1238;
Best Local Similarity 22.4%; Pred. No. 13;
Matches 55; Conservative 20; Mismatches 77; Indels 94; Gaps 11;

QY 2 GEAGV--PEKRPC-----EDLRC-----PETTS-----QALPAFTTEIQ 33
DB 947 GECGAEEPPSTPCLPBSGHLDDNNCARLTLHFNRDHVPQGTTVGALCSGIRSLPATRAVAR 1006
QY 34 E-----ASEGPGADEVOV--FAPANALPARS--EAAAVQPVIGISQVRNMSKEKD 81
DB 1007 DRLVLCDRASSGASAVEAVSFPARDLPDSSLICGAHAIVAATOR----- 1056
QY 82 LGTLGYVLGITMVTIITAIGAGIILG-----YSYRGKDLKE 118
DB 1057 -GNSSLLAVTEVKETVVTGSSSTGLVPLCGAFSVLMACVLCVWTRKRRK--- 1111
QY 119 QHDQKVCEREMQRT-----LPLSAFTNPTCEIVDEKTVVHTSQTVPDQEGSTP 169
DB 1112 -----ERERSRLPRESANNQWAPLNPRIERNRPGHKGDLVYQCKNFTPPRRADEA 1164
QY 170 LMGQAG 175
DB 1165 LGGPAG 1170

RESULT 10

US-09-855-722-5
; Sequence 5, Application US/09855722
; Patent No. US20020049306A1
; GENERAL INFORMATION:
; APPLICANT: Sakano, Sei-ji
; APPLICANT: Itoh, Akira
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-8576
; CURRENT APPLICATION NUMBER: US/09/855,722
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/214,278
; PRIOR FILING DATE: 1999-01-26
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1238
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-855-722-5

Query Match 8.7%; Score 80; DB 10; Length 1238;
Best Local Similarity 22.4%; Pred. No. 13;
Matches 55; Conservative 20; Mismatches 77; Indels 94; Gaps 11;

```
QY 2 GAGV--PEKRC-----EDLRC-----PETTS-----QALPAFTTEIQ 33
DB 947 GECGAEPPSTPCLPRSGHLNNCARLTLHFNRDHVPQGTIVGALCSGRSLPATRAVAR 1006
QY 34 E-----ASEPGADEVOV-FAPANALPARS--EAAAVQPVIGISORVRMSNEKKD 81
DB 1007 DELVLVLCDRASGASAVEVAFSPARDLPSSSLICGAHAIVAAITQR----- 1056
QY 82 LGTLGVGITMMVITIIAGIILG-----YXKRGDLKE 118
DB 1057 -GNSSLILAVTEVETVVGSSSTGLVPLVCGAFSVMLACVLCVWTRRRK--- 1111
QY 119 QHDQVCEREMOIT-----LPLSAFTNPCEIYDEKTVVHTSQTPVDPOGSTRP 169
DB 1112 -----ERERSRLPRESANMOMAPLNPRIPIRPGKHVDVLYQCKNTPPPRADEA 1164
QY 170 LMGQAG 175
DB 1165 LPPGAG 1170
```

```
RESULT 11
US-09-944-849-4
; Sequence 4, Application US/09944849
; Patent No. US2002015147A1
; GENERAL INFORMATION:
; APPLICANT: Nickoloff, Brian
; APPLICANT: Miele, Lucio
; TITLE OF INVENTION: METHOD AND REAGENTS FOR EPITHELIAL BARRIER FORMATION AND TREATMEN
; TITLE OF INVENTION: MALIGNANT AND BENIGN SKIN DISORDERS BY MODULATING THE NOTCH PATH
; FILE REFERENCE: 212583
; CURRENT APPLICATION NUMBER: US/09/944,849
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,614
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1238
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-944-849-4
```

```
Query Match 8.7%; Score 80; DB 10; Length 1238;
Best Local Similarity 22.4%; Pred. No. 13;
Matches 55; Conservative 20; Mismatches 77; Indels 94; Gaps 11;
```

```
QY 2 GAGV--PEKRC-----EDLRC-----PETTS-----QALPAFTTEIQ 33
DB 947 GECGAEPPSTPCLPRSGHLNNCARLTLHFNRDHVPQGTIVGALCSGRSLPATRAVAR 1006
QY 34 E-----ASEPGADEVOV-FAPANALPARS--EAAAVQPVIGISORVRMSNEKKD 81
DB 1007 DELVLVLCDRASGASAVEVAFSPARDLPSSSLICGAHAIVAAITQR----- 1056
QY 82 LGTLGVGITMMVITIIAGIILG-----YXKRGDLKE 118
DB 1057 -GNSSLILAVTEVETVVGSSSTGLVPLVCGAFSVMLACVLCVWTRRRK--- 1111
QY 119 QHDQVCEREMOIT-----LPLSAFTNPCEIYDEKTVVHTSQTPVDPOGSTRP 169
DB 1112 -----ERERSRLPRESANMOMAPLNPRIPIRPGKHVDVLYQCKNTPPPRADEA 1164
QY 170 LMGQAG 175
DB 1165 LPPGAG 1170
```

```
RESULT 12
US-09-921-667-6
; Sequence 6, Application US/09921667
; Patent No. US20020064527A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Mohler, Kendall M.
; APPLICANT: Barone, Dauphine S.
; APPLICANT: Peschon, Jacques J.
; APPLICANT: Kennedy, Mary K.
; APPLICANT: Plummer, John D.
; TITLE OF INVENTION: METHODS FOR TREATING AUTOIMMUNE AND CHRONIC INFLAMMATORY CONDITI
; TITLE OF INVENTION: ANTAGONISTS OF CD30 OR CD30L
; FILE REFERENCE: 2959-A
; CURRENT APPLICATION NUMBER: US/09/921,667
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/224,079
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 595
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-921-667-6
```

```
Query Match 8.5%; Score 77.5; DB 10; Length 595;
Best Local Similarity 23.6%; Pred. No. 8.9; Indels 43; Gaps 11;
Matches 46; Conservative 32; Mismatches 74;
```

```
QY 4 AGVEKRPCEIDRCPTTSQALPAFT-----TEIOASBGPGADEVOVFAPANALPARSEA 59
DB 307 AGETVTRK-QDMAEKDTFEAPLPQTDPDNCNTPENGCAFASST-----SPTOSILVDSQA 360
QY 60 AAQVPVIGISORVRMSNEKKDLGLGVVITMMVITIIAGIILGYSYKRGDLKEQ 119
DB 361 SKLTP-IPISAPVALSSTGKREVLDA-GPVLFWVILVIVVVGSSAFL-----L 406
QY 120 HDQVCEREMOIT-----LPLSAFTNPCEIYDEKTVVHTSQTPVDPOGSTRP 166
DB 407 CHRACKRIROKHLICPVOT-SQPKELVDSRP---RRSTQLRSASVTEPVAEKRG 462
QY 167 -STPLMGQAGTPGA 179
DB 463 LMSQPLMETCHSVGA 477
```

```
RESULT 13
US-09-870-759-51
; Sequence 51, Application US/09870759
; Patent No. US2002017551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; PRIOR FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51
; LENGTH: 1709
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-870-759-51
```

```
Query Match 8.5%; Score 77.5; DB 9; Length 1709;
```

```
Best Local Similarity 19.1%; Pred. No. 37; Indels 43; Gaps 6;
```

```
QY 22 SQALPAFTTEI-----QASBGPGADEVOVFAPANALPARSEA----- 59
DB 1558 SEPLASITLHLSGRVLAASSQPGAPAPEPHIVLASPNALRVDEIALRPSDQGEYICSAAN 1617
QY 60 ---AAQVPVIGISORVRMSNEKKDLGLGVVITMMVITIIAGIILGYSYKRGKD 115
DB 1618 VLGSASSTYFVGRALHRLHFOQ-----LWVLGLVGLVLLILILGAGC--YTWRR--- 1667
```

QY 116 LKEGHQKVCEREMORITLPLSAFTNPTCEIVDEKTVVHTS 157
 Db 1668 -----RRVCKQSMGENSENEM-APQKETTQIDPDATCETS 1702

RESULT 14
 US-09-748-537-14
 ; Sequence 14, Application US/09748537
 ; Patent No. US2002006183JA1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bertin, John
 ; APPLICANT: Chao, Moses V.
 ; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILIE AND USES THERE
 ; FILE REFERENCE: 07314-316001
 ; CURRENT APPLICATION NUMBER: US/09/748,537
 ; PRIOR FILING DATE: 2000-12-26
 ; PRIOR APPLICATION NUMBER: US 09/099,041
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: US 09/019,942
 ; PRIOR FILING DATE: 1998-02-06
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 14
 ; LENGTH: 425
 ; TYPE: PRT
 ; ORGANISM: Rattus rattus
 US-09-748-537-14

Query Match 8.4%; Score 76.5; DB 10; Length 425;
 Best Local Similarity 24.4%; Pred. No. 7.1; Indels 17; Gaps 5;
 Matches 42; Conservative 23; Mismatches

QY 3 EAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASGPGADEVQFAPANALPARSEAAAY 62
 Db 175 EROLRECTPWADACEEIPGRWIPRSTPEGSDSTAPSTQEPV-PEODLVSTVADNV 233
 QY 63 QPVIGISQVRMNSKEKKDGLTGLVIGITMMVIIAIGAGIILGYSGKRGDLKECHDQ 122
 Db 234 TTVMGSSQPVVTR-----GTTDNLIPV-YCSILAAVVGVLAVIAFKRMNSCK-QNKQ 284
 QY 123 KYCEREMORITLPLSAFTNPTCEIVDEKTVVHTSOTVPDPOBSTPLMGQA 174
 Db 285 GANSR-----PVMOTPPPEGKLSHDSGISVDSQSLHDQOTHTQTASGQA 329

RESULT 15
 US-09-821-831-2
 ; Sequence 2, Application US/09821831
 ; Patent No. US20020137188A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bartlett, Perry Francis
 ; APPLICANT: Coulson, Elizabeth Jane
 ; APPLICANT: Fieldew, Katrina
 ; APPLICANT: Baca, Manuel
 ; APPLICANT: Kilpatrick, Trevor
 ; APPLICANT: Surinder, Cheema
 ; TITLE OF INVENTION: Method of Modulating Cell Survival and
 ; FILE REFERENCE: 3206.1001-000
 ; CURRENT APPLICATION NUMBER: US/09/821,831
 ; PRIOR FILING DATE: 2001-03-30
 ; PRIOR APPLICATION NUMBER: PCT/AU99/00860
 ; PRIOR FILING DATE: 1999-10-05
 ; PRIOR APPLICATION NUMBER: AU PQ0702
 ; PRIOR FILING DATE: 1999-06-01
 ; PRIOR APPLICATION NUMBER: AU PP6351
 ; PRIOR FILING DATE: 1998-10-07
 ; PRIOR APPLICATION NUMBER: AU PP6353
 ; PRIOR FILING DATE: 1998-10-06
 ; NUMBER OF SEQ ID NOS: 72
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 425

; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Deduced amino acid sequence
 US-09-821-831-2

Query Match 8.4%; Score 76.5; DB 10; Length 425;
 Best Local Similarity 24.4%; Pred. No. 7.1; Indels 17; Gaps 5;
 Matches 42; Conservative 23; Mismatches
 QY 3 EAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASGPGADEVQFAPANALPARSEAAAY 62
 Db 175 EROLRECTPWADACEEIPGRWIPRSTPEGSDSTAPSTQEPV-PEODLVSTVADNV 233
 QY 63 QPVIGISQVRMNSKEKKDGLTGLVIGITMMVIIAIGAGIILGYSGKRGDLKECHDQ 122
 Db 234 TTVMGSSQPVVTR-----GTTDNLIPV-YCSILAAVVGVLAVIAFKRMNSCK-QNKQ 284
 QY 123 KYCEREMORITLPLSAFTNPTCEIVDEKTVVHTSOTVPDPOBSTPLMGQA 174
 Db 285 GANSR-----PVMOTPPPEGKLSHDSGISVDSQSLHDQOTHTQTASGQA 329

Search completed: April 7, 2003, 09:31:30
 Job time : 14.6528 secs

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:17:08 ; Search time 152.95 Seconds

(without alignments)
754.542 Million cell updates/sec

Title: US-10-057-951-2_COPY_85_263

Perfect score: 916
Sequence: 1 SGKACVPEKPCEDLRCPET.....PVDPEGSTPLMGQATPEA 179Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/PCRTUS_COMB.pep:*

2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*

3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*

4: /cgn2_6/ptodata/1/paa/US08_COMB.pep:*

5: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*

6: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*

7: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*

8: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*

9: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*

10: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*

11: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*

12: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*

13: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*

14: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*

15: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*

16: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*

17: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*

18: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*

19: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*

20: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*

21: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*

22: /cgn2_6/ptodata/1/paa/US099_COMB.pep:*

23: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*

24: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*

25: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*

26: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

27: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	916	100.0	250	14	US-09-001-403-64
2	916	100.0	263	1	Sequence 64, Appli
3	916	100.0	263	1	Sequence 2, Appli
4	916	100.0	263	1	Sequence 164, Appli
5	916	100.0	263	14	Sequence 2, Appli
6	916	100.0	263	15	Sequence 2, Appli

7	916	100.0	263	24	US-10-057-951-2	Sequence 2, Appli
8	916	100.0	263	25	US-10-102-704-2	Sequence 2, Appli
9	916	100.0	263	27	US-60-048-000-2	Sequence 2, Appli
10	913	99.7	257	1	PCT-US01-14827-8898	Sequence 8898, Ap
11	913	99.7	263	21	US-09-791-537-39562	Sequence 39562, A
12	913	99.7	263	23	US-09-927-796-44	Sequence 44, Appli
13	913	99.7	263	26	US-10-210-951-44	Sequence 44, Appli
14	913	99.7	263	26	US-10-211-858-44	Sequence 44, Appli
15	913	99.7	263	26	US-10-211-884-44	Sequence 44, Appli
16	913	99.7	263	27	US-60-230-435-1580	Sequence 1580, Ap
17	913	99.7	263	27	US-60-389-987-244	Sequence 244, App
18	913	99.7	263	27	US-60-412-418-244	Sequence 244, App
19	902	98.5	263	20	US-09-611-526-3727	Sequence 3727, Ap
20	851.5	93.0	293	1	PCT-US01-14827-8900	Sequence 8900, Ap
21	807	88.1	308	17	US-60-207-315-471	Sequence 471, App
22	704.5	76.9	286	15	US-09-144-889-4	Sequence 4, Appli
23	476	52.0	178	27	US-60-213-800-237	Sequence 237, App
24	426	46.5	88	22	US-09-834-366-16200	Sequence 16200, A
25	426	46.5	88	27	US-60-197-873-16200	Sequence 16200, A
26	322	35.2	66	1	PCT-US01-00663-31583	Sequence 31583, A
27	322	35.2	66	22	US-09-864-761-38457	Sequence 38457, A
28	322	35.2	66	25	US-10-182-993-30642	Sequence 30642, A
29	322	35.2	66	25	US-10-182-995-24929	Sequence 24929, A
30	322	35.2	66	25	US-10-182-997-23626	Sequence 23626, A
31	322	35.2	66	26	US-10-203-134-31343	Sequence 31343, A
32	322	35.2	66	26	US-10-203-135-30493	Sequence 30493, A
33	322	35.2	66	26	US-10-203-136-31400	Sequence 31400, A
34	322	35.2	66	26	US-10-203-137-31583	Sequence 31583, A
35	322	35.2	66	26	US-10-203-139-30540	Sequence 30540, A
36	275	30.0	146	1	PCT-US02-05301-237	Sequence 237, App
37	275	30.0	146	1	PCT-US02-05301-311	Sequence 311, App
38	223	24.3	126	17	US-09-307-140-979	Sequence 979, App
39	223	24.3	126	22	US-09-817-076-979	Sequence 1339, App
40	135	14.7	109	22	US-09-834-366-13839	Sequence 13839, A
41	135	14.7	109	27	US-60-197-873-18937	Sequence 8897, Ap
42	103	11.2	117	1	PCT-US01-14827-18937	Sequence 14827, A
43	94	10.3	4752	27	US-60-167-217-14372	Sequence 11554, A
44	94	10.3	4752	27	US-60-173-464-11654	Sequence 11654, A
45	94	10.3	4839	27	US-60-191-637-28375	Sequence 28375, A

ALIGNMENTS

RESULT 1

US-09-001-403-64

Sequence 64, Application US/09001403

GENERAL INFORMATION:

APPLICANT: Lai, Preeti

APPLICANT: Bandman, Olga

APPLICANT: Hillman, Jennifer L.

APPLICANT: Au-Young, Janice

APPLICANT: Tang, Y. Tom

APPLICANT: Yue, Henry

APPLICANT: Shah, Purvi

APPLICANT: Guegler, Karl J.

APPLICANT: Corley, Neil C.

TITLE OF INVENTION: HUMAN REGULATORY PROTEINS

NUMBER OF SEQUENCES: 150

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/001.403

```

FILING DATE: HERREWITH
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: BILINGS, LUCY J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0455 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 250 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: INODNOT05
CLONE: 3122252
US-09-001-403-64
```

```

Query Match          100.0%; Score 916; DB 14; Length 250;
Best Local Similarity 100.0%; Pred. No. 8.6e-90;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 SGEAGVPEKRPCEDLCPEPTTSQALPAFTTEIOEASGEGADEVQVFAPANLPAARSEAA 60
DB 72 SGEAGVPEKRPCEDLCPEPTTSQALPAFTTEIOEASGEGADEVQVFAPANLPAARSEAA 131
QY 61 AVQPVIGISQVRMNSKEKKDLGTLGYVLGITMVIITAIAGAILIGYSYKRGKDLKEQH 120
DB 132 AVQPVIGISQVRMNSKEKKDLGTLGYVLGITMVIITAIAGAILIGYSYKRGKDLKEQH 191
QY 121 DQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTSQTPVDQEGSTPLMGAGTPGA 179
DB 192 DQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTSQTPVDQEGSTPLMGAGTPGA 250
```

```

RESULT 2
PCT-US00-27239-2
; Sequence 2, Application PC/TUS0027239
; GENERAL INFORMATION:
; APPLICANT: HUMAN GENOME SCIENCES, INC.
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
; FILE REFERENCE: PF378PCT2
; CURRENT APPLICATION NUMBER: PCT/US00/27239
; CURRENT FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 09/411,977
; PRIOR FILING DATE: 1999-10-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US00-27239-2
```

```

Query Match          100.0%; Score 916; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 9.3e-90;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 SGEAGVPEKRPCEDLCPEPTTSQALPAFTTEIOEASGEGADEVQVFAPANLPAARSEAA 60
DB 85 SGEAGVPEKRPCEDLCPEPTTSQALPAFTTEIOEASGEGADEVQVFAPANLPAARSEAA 144
QY 61 AVQPVIGISQVRMNSKEKKDLGTLGYVLGITMVIITAIAGAILIGYSYKRGKDLKEQH 120
DB 145 AVQPVIGISQVRMNSKEKKDLGTLGYVLGITMVIITAIAGAILIGYSYKRGKDLKEQH 204
QY 121 DQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTSQTPVDQEGSTPLMGAGTPGA 179
DB 205 DQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTSQTPVDQEGSTPLMGAGTPGA 263
```

```

RESULT 3
PCT-US02-05301-164
; Sequence 164, Application PC/TUS0205301
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: PS736PCT
; CURRENT APPLICATION NUMBER: PCT/US02/05301
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/304,417
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US 60/270,625
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 164
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-05301-164
```

```

Query Match          100.0%; Score 916; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 9.3e-90;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 SGEAGVPEKRPCEDLCPEPTTSQALPAFTTEIOEASGEGADEVQVFAPANLPAARSEAA 60
DB 85 SGEAGVPEKRPCEDLCPEPTTSQALPAFTTEIOEASGEGADEVQVFAPANLPAARSEAA 144
QY 61 AVQPVIGISQVRMNSKEKKDLGTLGYVLGITMVIITAIAGAILIGYSYKRGKDLKEQH 120
DB 145 AVQPVIGISQVRMNSKEKKDLGTLGYVLGITMVIITAIAGAILIGYSYKRGKDLKEQH 204
QY 121 DQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTSQTPVDQEGSTPLMGAGTPGA 179
DB 205 DQKVCEREMQRIITPLSAFTNPCEIYDEKTVVHTSQTPVDQEGSTPLMGAGTPGA 263
```

```

RESULT 4
PCT-US98-10728-2
; Sequence 2, Application PC/TUS9810728
; GENERAL INFORMATION:
; APPLICANT: EBNER, REINHARD
; APPLICANT: MOORE, PAUL
; APPLICANT: RUBEN, STEVEN
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/10728
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKS, ANDERS A.
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF379PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8439
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
```


TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US98-10728-2

Query Match 100.0%; Score 916; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 9.3e-90;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIOEASGPGADEVQVAPANALPARSEAA 60
DB 85 SGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIOEASGPGADEVQVAPANALPARSEAA 144
QY 61 AVQPVIGISQVRMNSKEKDLGTLGVLGITMMVILIIAGAGIILIGYSYKRGDKLEQH 120
DB 145 AVQPVIGISQVRMNSKEKDLGTLGVLGITMMVILIIAGAGIILIGYSYKRGDKLEQH 204
QY 121 DQKVCEREMORITLPLSAFTNPCEIYDEKTVVHTSQTVPDQEGSTPLMGAGTRGA 179
DB 205 DQKVCEREMORITLPLSAFTNPCEIYDEKTVVHTSQTVPDQEGSTPLMGAGTRGA 263

RESULT 5

US-09-084-491A-2
Sequence 2, Application US/09084491A
GENERAL INFORMATION:
APPLICANT: MOORE, PAUL A.
APPLICANT: RUBEN, STEVEN M.
APPLICANT: BERNER, REINHARD
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/084,491A
FILING DATE: 27-MAY-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF378
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-084-491A-2

Query Match 100.0%; Score 916; DB 14; Length 263;
Best Local Similarity 100.0%; Pred. No. 9.3e-90;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIOEASGPGADEVQVAPANALPARSEAA 60
DB 85 SGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIOEASGPGADEVQVAPANALPARSEAA 144
QY 61 AVQPVIGISQVRMNSKEKDLGTLGVLGITMMVILIIAGAGIILIGYSYKRGDKLEQH 120
DB 145 AVQPVIGISQVRMNSKEKDLGTLGVLGITMMVILIIAGAGIILIGYSYKRGDKLEQH 204

QY 121 DQKVCEREMORITLPLSAFTNPCEIYDEKTVVHTSQTVPDQEGSTPLMGAGTRGA 179
DB 205 DQKVCEREMORITLPLSAFTNPCEIYDEKTVVHTSQTVPDQEGSTPLMGAGTRGA 263

RESULT 6

US-09-144-889-2
Sequence 2, Application US/09144889B
GENERAL INFORMATION:
APPLICANT: Earl F. Albore
APPLICANT: Kristine K. Kikly
TITLE OF INVENTION: KRINGLE1
FILE REFERENCE: GH-70249
CURRENT APPLICATION NUMBER: US/09/144,889B
CURRENT FILING DATE: 1998-09-01
EARLIER APPLICATION NUMBER: 60/056,032
EARLIER FILING DATE: 1997-09-02
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 263
TYPE: PRT
ORGANISM: HOMO SAPIENS
US-09-144-889-2

Query Match 100.0%; Score 916; DB 15; Length 263;
Best Local Similarity 100.0%; Pred. No. 9.3e-90;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIOEASGPGADEVQVAPANALPARSEAA 60
DB 85 SGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIOEASGPGADEVQVAPANALPARSEAA 144
QY 61 AVQPVIGISQVRMNSKEKDLGTLGVLGITMMVILIIAGAGIILIGYSYKRGDKLEQH 120
DB 145 AVQPVIGISQVRMNSKEKDLGTLGVLGITMMVILIIAGAGIILIGYSYKRGDKLEQH 204
QY 121 DQKVCEREMORITLPLSAFTNPCEIYDEKTVVHTSQTVPDQEGSTPLMGAGTRGA 179
DB 205 DQKVCEREMORITLPLSAFTNPCEIYDEKTVVHTSQTVPDQEGSTPLMGAGTRGA 263

RESULT 7

US-10-057-951-2
Sequence 2, Application US/10057951
GENERAL INFORMATION:
APPLICANT: Moore et al.
TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
FILE REFERENCE: PF378P1
CURRENT APPLICATION NUMBER: US/10/057,951
CURRENT FILING DATE: 2002-01-29
PRIOR APPLICATION NUMBER: US 09/411,977
PRIOR FILING DATE: 1999-10-04
PRIOR APPLICATION NUMBER: US 09/084,491
PRIOR FILING DATE: 1998-05-27
PRIOR APPLICATION NUMBER: US 60/048,000
PRIOR FILING DATE: 1997-05-28
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 2
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
US-10-057-951-2

Query Match 100.0%; Score 916; DB 24; Length 263;
Best Local Similarity 100.0%; Pred. No. 9.3e-90;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIOEASGPGADEVQVAPANALPARSEAA 60
DB 85 SGEAGVPEKRPCEDLRCPEFTTSQALPAFTTEIOEASGPGADEVQVAPANALPARSEAA 144

QY 61 AVQPVIGISQVRMNSKEKKDLGTLGYVLGTTMVTIIAIGAGIILGYSKRGDKLEQH 120
DB 145 AVQPVIGISQVRMNSKEKKDLGTLGYVLGTTMVTIIAIGAGIILGYSKRGDKLEQH 204
QY 121 DQKVCEREMORITLPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGQAGTPGA 179
DB 205 DQKVCEREMORITLPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGQAGTPGA 263

RESULT 8
US-10-102-704-2
Sequence 2, Application US/10102704
GENERAL INFORMATION:

APPLICANT: Moore et al.
TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protein
FILE REFERENCE: PF378CI
CURRENT APPLICATION NUMBER: US/10/102,704
CURRENT FILING DATE: 2002-03-22
PRIOR APPLICATION NUMBER: 09/084,491
PRIOR FILING DATE: 1998-05-27
PRIOR APPLICATION NUMBER: 60/048,000
PRIOR FILING DATE: 1997-05-28
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin version 3.1
SEQ ID NO 2
LENGTH: 263
TYPE: PRT
ORGANISM: Homo sapiens
US-10-102-704-2

Query Match 100.0%; Score 916; DB 25; Length 263;
Best Local Similarity 100.0%; Pred. No. 9.3e-90;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGEAGVPEKRCPCEDLRCPETTSQALPAFTTEIOEASGPGADEVQVAPANALPARSEAA 60
DB 85 SGEAGVPEKRCPCEDLRCPETTSQALPAFTTEIOEASGPGADEVQVAPANALPARSEAA 144
QY 61 AVQPVIGISQVRMNSKEKKDLGTLGYVLGTTMVTIIAIGAGIILGYSKRGDKLEQH 120
DB 145 AVQPVIGISQVRMNSKEKKDLGTLGYVLGTTMVTIIAIGAGIILGYSKRGDKLEQH 204
QY 121 DQKVCEREMORITLPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGQAGTPGA 179
DB 205 DQKVCEREMORITLPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGQAGTPGA 263

RESULT 9

US-60-048-000-2
Sequence 2, Application US/60048000
GENERAL INFORMATION:

APPLICANT: EBNER, REINHARD
APPLICANT: MOORE, PAUL
APPLICANT: RUBEN, STEVE
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/60/048,000
FILING DATE:

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF379PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-60-048-000-2

Query Match 100.0%; Score 916; DB 27; Length 263;
Best Local Similarity 100.0%; Pred. No. 9.3e-90;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGEAGVPEKRCPCEDLRCPETTSQALPAFTTEIOEASGPGADEVQVAPANALPARSEAA 60
DB 85 SGEAGVPEKRCPCEDLRCPETTSQALPAFTTEIOEASGPGADEVQVAPANALPARSEAA 144
QY 61 AVQPVIGISQVRMNSKEKKDLGTLGYVLGTTMVTIIAIGAGIILGYSKRGDKLEQH 120
DB 145 AVQPVIGISQVRMNSKEKKDLGTLGYVLGTTMVTIIAIGAGIILGYSKRGDKLEQH 204
QY 121 DQKVCEREMORITLPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGQAGTPGA 179
DB 205 DQKVCEREMORITLPLSAFTNPCEIYDEKTVVHTSQTPVDPOEGSTPLMGQAGTPGA 263

RESULT 10
PCT-US01-14827-8898
Sequence 8898, Application PC/TUS0114827
GENERAL INFORMATION:

APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-104
CURRENT APPLICATION NUMBER: PCT/US01/14827
CURRENT FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: 09/577,408
PRIOR FILING DATE: 2000-05-18
NUMBER OF SEQ ID NOS: 16102
SOFTWARE: Custom
SEQ ID NO 8898
LENGTH: 257
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (57)-(78)
OTHER INFORMATION: KRINGLE DOMAIN SIGNATURE domain identified by eMATRIX,
OTHER INFORMATION: accession number PR00018C, p-value=5.235e-13, raw score of 14.30
LOCATION: (125)-(196)
OTHER INFORMATION: kringle domain identified by Pfam, accession name kringle_E-
OTHER INFORMATION: value=1.9e-05, Pfam score of 10.1
PCT-US01-14827-8898

Query Match 99.7%; Score 913; DB 1; Length 257;
Best Local Similarity 99.4%; Pred. No. 1.9e-89;
Matches 178; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGEAGVPEKRCPCEDLRCPETTSQALPAFTTEIOEASGPGADEVQVAPANALPARSEAA 60
DB 79 SGEAGVPEKRCPCEDLRCPETTSQALPAFTTEIOEASGPGADEVQVAPANALPARSEAA 138
QY 61 AVQPVIGISQVRMNSKEKKDLGTLGYVLGTTMVTIIAIGAGIILGYSKRGDKLEQH 120
DB 139 AVQPVIGISQVRMNSKEKKDLGTLGYVLGTTMVTIIAIGAGIILGYSKRGDKLEQH 198

```

QY      121 DOKVDEREMORTLPLSAFTNPTCEIVDEKTVVHTSQTVPVDPOEGSTPLMGAGTTPGA 179
          |||||||
DB      199 DOKVDEREMORTLPLSAFTNPTCEIVDEKTVVHTSQTVPVDPOEGSTPLMGAGTTPGA 257

RESULT 11
US-09-791-537-39562
; Sequence 39562, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 39562
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-39562

Query Match           99.7%; Score 913; DB 21; Length 263;
Best Local Similarity 99.4%; Pred. No. 2e-89;
Matches 178; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 SEAGVPERPRCPEDLRCPETTSOALPAFTTEIOEASEGGADEVGFAPANALPARSEAA 60
          |||||
DB      85 SEAGVPERPRCPEDLRCPETTSOALPAFTTEIOEASEGGADEVGFAPANALPARSEAA 144

QY      61 AVQPVIIGISQRIRMSKSKKDGLTGVLGITMMVTIIIAIGAILLGYSYRKGDKLEQH 120
          |||||
DB      145 AVQPVIIGISQRIRMSKSKKDGLTGVLGITMMVTIIIAIGAILLGYSYRKGDKLEQH 204

QY      121 DOKVDEREMORTLPLSAFTNPTCEIVDEKTVVHTSQTVPVDPOEGSTPLMGAGTTPGA 179
          |||||||
DB      205 DOKVDEREMORTLPLSAFTNPTCEIVDEKTVVHTSQTVPVDPOEGSTPLMGAGTTPGA 263

RESULT 12
US-09-927-796-44
; Sequence 44, Application US/09927796
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austen L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Masters, Scot A.
; APPLICANT: Pan, James
; APPLICANT: Pittl, Robert M.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stone, Donna M.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
; FILE REFERENCE: P2931RCl
; CURRENT APPLICATION NUMBER: US/09/927,796
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/014699
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 60/026943
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 60/059121
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/062037
; PRIOR FILING DATE: 1997-10-10

```

;; PRIOR FILING DATE: 1999-04-30
;; PRIOR APPLICATION NUMBER: 09/380137
;; PRIOR FILING DATE: 1999-08-25
;; PRIOR APPLICATION NUMBER: 09/380138
;; PRIOR FILING DATE: 1999-08-25
;; PRIOR APPLICATION NUMBER: 09/380139
;; PRIOR FILING DATE: 1999-08-25
;; PRIOR APPLICATION NUMBER: 09/403297
;; PRIOR FILING DATE: 1999-10-18
;; PRIOR APPLICATION NUMBER: 09/423844
;; PRIOR FILING DATE: 1999-11-12
;; PRIOR APPLICATION NUMBER: 09/511133
;; PRIOR FILING DATE: 2000-02-23
;; PRIOR APPLICATION NUMBER: 09/511631
;; PRIOR FILING DATE: 2000-02-23
;; PRIOR APPLICATION NUMBER: 09/664610
;; PRIOR FILING DATE: 2000-09-18
;; PRIOR APPLICATION NUMBER: 09/665350
;; PRIOR FILING DATE: 2000-09-18
;; PRIOR APPLICATION NUMBER: 09/690169
;; PRIOR FILING DATE: 2000-10-16
;; PRIOR APPLICATION NUMBER: 09/690189
;; PRIOR FILING DATE: 2000-10-16
;; PRIOR APPLICATION NUMBER: 09/709238
;; PRIOR FILING DATE: 2000-11-18
;; PRIOR APPLICATION NUMBER: 09/866034
;; PRIOR FILING DATE: 2001-05-25
;; PRIOR APPLICATION NUMBER: 09/872035
;; PRIOR FILING DATE: 2001-06-01
;; PRIOR APPLICATION NUMBER: 09/884733
;; PRIOR FILING DATE: 2001-06-19
;; PRIOR APPLICATION NUMBER: 09/886342
;; PRIOR FILING DATE: 2001-06-19
;; PRIOR APPLICATION NUMBER: 09/866028
;; PRIOR FILING DATE: 2001-08-25
;; PRIOR APPLICATION NUMBER: PCT/US97/05230
;; PRIOR FILING DATE: 1997-03-31
;; PRIOR APPLICATION NUMBER: PCT/US98/19094
;; PRIOR FILING DATE: 1998-09-14
;; PRIOR APPLICATION NUMBER: PCT/US98/19330
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: PCT/US98/21407
;; PRIOR FILING DATE: 1998-10-09
;; PRIOR APPLICATION NUMBER: PCT/US98/25108
;; PRIOR FILING DATE: 1998-12-01
;; PRIOR APPLICATION NUMBER: PCT/US99/05028
;; PRIOR FILING DATE: 1999-03-08
;; PRIOR APPLICATION NUMBER: PCT/US99/12252
;; PRIOR FILING DATE: 1999-06-02
;; PRIOR APPLICATION NUMBER: PCT/US99/20111
;; PRIOR FILING DATE: 1999-09-01
;; PRIOR APPLICATION NUMBER: PCT/US99/21090
;; PRIOR FILING DATE: 1999-09-15
;; PRIOR APPLICATION NUMBER: PCT/US99/28313
;; PRIOR FILING DATE: 1999-11-30
;; PRIOR APPLICATION NUMBER: PCT/US99/28301
;; PRIOR FILING DATE: 1999-12-01
;; PRIOR APPLICATION NUMBER: PCT/US99/28634
;; PRIOR FILING DATE: 1999-12-01
;; PRIOR APPLICATION NUMBER: PCT/US99/28565
;; PRIOR FILING DATE: 1999-12-02
;; PRIOR APPLICATION NUMBER: PCT/US00/00219
;; PRIOR FILING DATE: 2000-01-05
;; PRIOR APPLICATION NUMBER: PCT/US00/03565
;; PRIOR FILING DATE: 2000-02-11
;; PRIOR APPLICATION NUMBER: PCT/US00/04341
;; PRIOR FILING DATE: 2000-02-18
;; PRIOR APPLICATION NUMBER: PCT/US00/04342
;; PRIOR FILING DATE: 2000-02-18
;; PRIOR APPLICATION NUMBER: PCT/US00/04414
;; PRIOR FILING DATE: 2000-02-22
;; PRIOR APPLICATION NUMBER: PCT/US00/05841
;; PRIOR FILING DATE: 2000-03-02

;; PRIOR APPLICATION NUMBER: PCT/US00/06884
;; PRIOR FILING DATE: 2000-03-15
;; PRIOR APPLICATION NUMBER: PCT/US00/08439
;; PRIOR FILING DATE: 2000-03-30
;; PRIOR APPLICATION NUMBER: PCT/US00/13705
;; PRIOR FILING DATE: 2000-05-17
;; PRIOR APPLICATION NUMBER: PCT/US00/14941
;; PRIOR FILING DATE: 2000-05-30
;; PRIOR APPLICATION NUMBER: PCT/US00/15264
;; PRIOR FILING DATE: 2000-06-02
;; PRIOR APPLICATION NUMBER: PCT/US00/20710
;; PRIOR FILING DATE: 2000-07-28
;; PRIOR APPLICATION NUMBER: PCT/US00/32678
;; PRIOR FILING DATE: 2000-12-01
;; PRIOR APPLICATION NUMBER: PCT/US01/17800
;; PRIOR FILING DATE: 2001-06-01
;; PRIOR APPLICATION NUMBER: PCT/US01/19692
;; PRIOR FILING DATE: 2001-06-20
;; PRIOR APPLICATION NUMBER: PCT/US01/21066
;; PRIOR FILING DATE: 2001-06-29
;; PRIOR APPLICATION NUMBER: PCT/US01/21735
;; PRIOR FILING DATE: 2001-07-09
;; NUMBER OF SEQ ID NOS: 258
;; SEQ ID NO 44
;; LENGTH: 263

Query Match 99.7%; Score 913; DB 23; Length 263;
Best Local Similarity 99.4%; Pred. No. 2e-89; 0; Indels 0; Gaps 0;
Matches 178; Conservative 1; Mismatches 0;

QY 1 SGEAGVPEKRCEDLRCPETTSQALPAFTTEIQEASGSGPADEVQVAPANALPARSEAA 60
DB 85 SGEAGVPEKRCEDLRCPETTSQALPAFTTEIQEASGSGPADEVQVAPANALPARSEAA 144

QY 61 AVQPVIGISQVRNNSKKDLGTLGYLGTVMVVIILAGITLLGYSGKGDLEQH 120
DB 145 AVQPVIGISQVRNNSKKDLGTLGYLGTVMVVIILAGITLLGYSGKGDLEQH 204

QY 121 DQKVCEREMORITLPLSAFTNPTCEIVDEKTVVHTSOTPDPOGSGTPLMGQGTGGA 179
DB 205 DQKVCEREMORITLPLSAFTNPTCEIVDEKTVVHTSOTPDPOGSGTPLMGQGTGGA 263

RESULT 13
US-10-210-951-44
;; Sequence 44, Application US/10210951
;; GENERAL INFORMATION:
;; APPLICANT: Ashkenazi, Avi J.
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Hillan, Kenneth J.
;; APPLICANT: Marsters, Scot A.
;; APPLICANT: Pan, James
;; APPLICANT: Pilti, Robert M.
;; APPLICANT: Roy, Margaret Ann
;; APPLICANT: Smith, Victoria
;; APPLICANT: Stone, Donna M.
;; APPLICANT: Watanebe, Colin K.
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
;; FILE REFERENCE: P2931R1C1
;; CURRENT APPLICATION NUMBER: US/10/210,951
;; CURRENT FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: 60/014699
;; PRIOR FILING DATE: 1996-04-01
;; PRIOR APPLICATION NUMBER: 60/026943
;; PRIOR FILING DATE: 1996-09-23
;; PRIOR APPLICATION NUMBER: 60/059121
;; PRIOR FILING DATE: 1997-07-17
;; PRIOR APPLICATION NUMBER: 60/059352
;; PRIOR FILING DATE: 1997-09-19
;; PRIOR APPLICATION NUMBER: 60/062037

```
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 258
; SEQ ID NO 44
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-210-951-44
```

```
Query Match          99.7%; Score 913; DB 26; Length 263;
Best Local Similarity 99.4%; Pred. No. 2e-89; 0; Indels 0; Gaps 0;
Matches 178; Conservative 1; Mismatches 0;
```

```
Qy 1 SGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASEGGADEVQVFAPANALPARSEAA 60
Db 85 SGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASEGGADEVQVFAPANALPARSEAA 144
Qy 61 AVQPVIGISQVRVMSKEKDLGTLGYVLGITMVIITIAIGIILGYSYKRGKDLKEQH 120
Db 145 AVQPVIGISQVRVMSKEKDLGTLGYVLGITMVIITIAIGIILGYSYKRGKDLKEQH 204
Qy 121 DQKVCEREMQRTITPLSAFTNPTCEIVDEKTVVHTSQTVPDPQEGSTPLMGAGTPGA 179
Db 205 DQKVCEREMQRTITPLSAFTNPTCEIVDEKTVVHTSQTVPDPQEGSTPLMGAGTPGA 263
```

```
RESULT 14
US-10-211-858-44
; Sequence 44, Application US/10211858
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Goddard, Audrey
; APPLICANT: Goddard, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Marsters, Scott A.
; APPLICANT: Pan, James
; APPLICANT: Pitti, Robert M.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stone, Donna M.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
; FILE REFERENCE: P2931R1C1
; CURRENT APPLICATION NUMBER: US/10/211,858
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/014699
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 60/026943
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 60/059121
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/062037
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
```

```
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 258
; SEQ ID NO 44
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-858-44
```

```
Query Match          99.7%; Score 913; DB 26; Length 263;
Best Local Similarity 99.4%; Pred. No. 2e-89; 0; Indels 0; Gaps 0;
Matches 178; Conservative 1; Mismatches 0;
```

```
Qy 1 SGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASEGGADEVQVFAPANALPARSEAA 60
Db 85 SGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASEGGADEVQVFAPANALPARSEAA 144
Qy 61 AVQPVIGISQVRVMSKEKDLGTLGYVLGITMVIITIAIGIILGYSYKRGKDLKEQH 120
Db 145 AVQPVIGISQVRVMSKEKDLGTLGYVLGITMVIITIAIGIILGYSYKRGKDLKEQH 204
Qy 121 DQKVCEREMQRTITPLSAFTNPTCEIVDEKTVVHTSQTVPDPQEGSTPLMGAGTPGA 179
Db 205 DQKVCEREMQRTITPLSAFTNPTCEIVDEKTVVHTSQTVPDPQEGSTPLMGAGTPGA 263
```

```
RESULT 15
US-10-211-884-44
; Sequence 44, Application US/10211884
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Goddard, Audrey
; APPLICANT: Goddard, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Marsters, Scott A.
; APPLICANT: Pan, James
; APPLICANT: Pitti, Robert M.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stone, Donna M.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF TUMOR
; FILE REFERENCE: P2931R1C1
; CURRENT APPLICATION NUMBER: US/10/211,884
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/014699
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: 60/026943
; PRIOR FILING DATE: 1996-09-23
; PRIOR APPLICATION NUMBER: 60/059121
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/062037
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 258
; SEQ ID NO 44
```

; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-884-44

Query Match 99.7%; Score 913; DB 26; Length 263;
Best Local Similarity 99.4%; Pred No. 2e-89; 0; Indels 0; Gaps 0;
Matches 178; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASEGPGADEVOVFAPANALPARSEEA 60
Db 85 SGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASEGPGADEVOVFAPANALPARSEEA 144
Qy 61 AVQPVIGISORVAMNSKEKKDLGTLGVVIGITMVTIIAIGAGIILGYSKRGKDLKEQH 120
Db 145 AVQPVIGISORVAMNSKEKKDLGTLGVVIGITMVTIIAIGAGIILGYSKRGKDLKEQH 204
Qy 121 DQKVCEREMORTLPLSAFTNPTCEIYDEKTVVHTSQTPVDPOEGSTPLMGQAGTPGA 179
Db 205 DQKVCEREMORTLPLSAFTNPTCEIYDEKTVVHTSQTPVDPOEGSTPLMGQAGTPGA 263

Search completed: April 7, 2003, 09:29:28
Job time : 153.95 secs

GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:18:38 ; Search time 26.7942 Seconds
(without alignments)
1010.317 Million cell updates/sec

Title: US-10-057-951-2_COPY_85_263

Perfect score: 916
Sequence: 1 SGEAGVPEKRPCEDLRCPEPT...PVDPOGSTRPLMGQAGTGA 179

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 679521 seqs, 151232488 residues

679521

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New.*

1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*

2: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*

3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*

4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*

5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*

6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*

7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	913	99.7	263	6	US-10-218-140-6002 Sequence 6002, App
2	732.5	80.0	264	6	US-10-144-779-488 Sequence 488, App
3	95	10.4	163	1	PCT-US02-332727-117039 Sequence 17039, A
4	95	10.4	163	5	US-09-978-825-17039 Sequence 17039, A
5	95	10.4	163	6	US-10-057-998-17039 Sequence 17039, A
6	87.5	9.6	850	6	US-10-369-493-10111 Sequence 10111, A
7	86.5	9.4	507	6	US-10-289-762-174 Sequence 174, App
8	86.5	9.4	820	6	US-10-282-122A-54852 Sequence 54852, A
9	80.5	8.8	769	5	US-09-701-271A-4 Sequence 4, Appli
10	80	8.7	345	5	US-09-724-676-63523 Sequence 63523, A
11	80	8.7	345	5	US-09-724-676A-63523 Sequence 63523, A
12	80	8.7	1226	5	US-09-724-676-63525 Sequence 63525, A
13	80	8.7	1226	5	US-09-724-676A-63525 Sequence 63525, A
14	80	8.7	1264	5	US-09-724-676-63524 Sequence 63524, A
15	80	8.7	1264	5	US-09-724-676A-63524 Sequence 63524, A
16	79.5	8.7	769	5	US-09-701-271A-6 Sequence 6, Appli
17	79.5	8.7	769	6	US-10-320-800-10 Sequence 10, Appli
18	78.5	8.6	2824	6	US-10-369-493-5166 Sequence 5166, App
19	77.5	8.5	452	6	US-10-282-122A-70552 Sequence 70552, A
20	77.5	8.5	463	6	US-10-092-411A-3973 Sequence 3973, App
21	77.5	8.5	977	7	US-60-440-068-526 Sequence 526, App
22	77.5	8.5	1709	5	US-09-751-708A-51 Sequence 51, Appli
23	77.5	8.5	1709	5	US-09-949-016-10503 Sequence 10503, A
24	77.5	8.5	1709	7	US-60-440-068-168 Sequence 168, App
25	77.5	8.5	1709	7	US-60-453-135-14575 Sequence 14575, A
26	77.5	8.5	1709	7	US-60-453-050-14575 Sequence 14575, A

ALIGNMENTS

27	77	8.4	769	5	US-09-701-271A-2	Sequence 2, Appli
28	76	8.3	490	6	US-10-369-493-21753	Sequence 21753, A
29	76	8.3	755	6	US-10-156-761-11835	Sequence 11835, A
30	75.5	8.2	219	6	US-10-092-411A-5651	Sequence 5651, App
31	75.5	8.2	276	6	US-10-203-138A-14692	Sequence 14692, A
32	75.5	8.2	422	1	PCT-US02-24310-1134	Sequence 134, App
33	75	8.2	595	5	US-09-949-016-8048	Sequence 6048, App
34	75	8.2	595	5	US-09-826-212A-9	Sequence 9, Appli
35	75	8.2	595	7	US-60-443-566-4128	Sequence 4128, App
36	75	8.2	595	7	US-60-452-680-20991	Sequence 20991, A
37	75	8.2	595	7	US-60-453-135-12718	Sequence 12718, A
38	75	8.2	595	7	US-60-453-050-12718	Sequence 12718, A
39	75	8.2	642	5	US-09-949-016-8043	Sequence 8043, App
40	75	8.2	774	5	US-09-905-846A-69	Sequence 69, Appli
41	75	8.2	774	5	US-09-948-429C-69	Sequence 69, Appli
42	75	8.2	788	6	US-10-156-761-14497	Sequence 14497, A
43	75	8.2	950	6	US-10-366-683-29012	Sequence 29012, A
44	74.5	8.1	340	6	US-10-282-122A-48299	Sequence 48299, A
45	74.5	8.1	630	7	US-60-452-680-23441	Sequence 23441, A

```
RESULT 1
US-10-218-140-6002
; Sequence 6002, Application US/10218140
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYPEPTIDES
; FILE REFERENCE: 15966-543 CON
; CURRENT APPLICATION NUMBER: US/01218, 140
; PRIOR FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 09/540, 763
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/127, 728
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/127, 636
; PRIOR FILING DATE: 1999-04-02
; PRIOR APPLICATION NUMBER: 60/127, 607
; NUMBER OF SEQ ID NOS: 6322
; SOFTWARE: Curator Version 1.0
; SEQ ID NO 6002
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-218-140-6002
Query Match 99.7%; Score 913; DB 6; Length 263;
Best Local Similarity 99.4%; Pred. No. 7.3e-81;
Matches 178; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SGEAGVPEKRPCEDLRCPEPTSOALPAFTTRIOASBEGADVEQVAPANALPARSEAA 60
DB 85 SGEAGVPEKRPCEDLRCPEPTSOALPAFTTRIOASBEGADVEQVAPANALPARSEAA 144
QY 61 AVOPVIGISQRYVMSKSKKDLGTLGYVIGITVMYIIAIGAGIILGYSKGKLEKH 120
DB 145 AVOPVIGISQRYVMSKSKKDLGTLGYVIGITVMYIIAIGAGIILGYSKGKLEKH 204
QY 121 DQVVCEREMORTLPLSAFTTFCIVDEKVVVHTSGTPVDPOGSTRPLMGQAGTGA 179
DB 205 DQVVCEREMORTLPLSAFTTFCIVDEKVVVHTSGTPVDPOGSTRPLMGQAGTGA 263
RESULT 2
US-10-144-779-488
; Sequence 488, Application US/10144779
; GENERAL INFORMATION:
; APPLICANT: SUBRAMANIAN, Mani et al.
```

;; TITLE OF INVENTION: MOUSE ORTHOLOGS OF HUMAN DISEASE GENES,
;; TITLE OF INVENTION: PROTEINS ENCODED BY THESE MOUSE GENES, AND USES THEREOF
;; FILE REFERENCE: CL001235
;; CURRENT APPLICATION NUMBER: US/10/144,779
;; CURRENT FILING DATE: 2002-05-15
;; NUMBER OF SEQ ID NOS: 864
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 488
;; LENGTH: 264
;; TYPE: PRT
;; ORGANISM: Mus musculus
US-10-144-779-488

Query Match 80.4%; Score 732.5; DB 6; Length 264;
Best Local Similarity 79.4%; Pred. No. 2.8e-63;
Matches 143; Conservative 13; Mismatches 23; Indels 1; Gaps 1;

QY 1 SGEAGVPEKRPCEDLCPERTSQA-LPATTIEQASSEGADDEVQVAPNALPARESEA 59
DB 85 SSETGVPEKRPCEDLCPERTSQA-LPATTIEQASSEGADDEVQVAPNALPARESEA 144
QY 60 AAVQPVIGISQVRNMSKEKDLGTLGYVLGITMVIITIAIGAGIIGSYKRGKDLKEQ 119
DB 145 AAVQPVIGISQVRNMSKEKDLGTLGYVLGITMVIITIAIGAGIIGSYKRGKDLKEQ 204
QY 120 HDQKVCEREMORTPLSFTNPTCEIVDEKVVVHTSQTVPDPOGSLPMGAKTPEGA 179
DB 205 HEKACEREMORTPLSFTNPTCEIVDEKVVVHTSQTVPDPOGSLPMGAKTPEGA 264

RESULT 3

PCT-US02-32727-17039
; Sequence 17039, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Derrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglas, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 17039
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Propionibacterium acnes
PCT-US02-32727-17039

Query Match 10.4%; Score 95; DB 1; Length 163;
Best Local Similarity 26.1%; Pred. No. 0.18;
Matches 40; Conservative 30; Mismatches 47; Indels 36; Gaps 9;

QY 49 PANALPARESEAAVQPVIGISQVRNMSKEKDLGTLGYVLGITMVIITIAIGAGIITL 106
DB 18 PQQARTAAARPTSLRPVLRALPQVR-----QQQWGRLGFAI--LIVVLAAGLAGLIV 69
QY 107 -----GYSYKRGKDLKE-----QHDQKVCEREMORTPL-----LSAFTNPTCEIV 147
DB 78 LNTTIOAQSMQIAQETRNINLVLQHQAVALAABVDHLRGQNLQEQAKKLGMRPNPYGSYI 129
QY 148 DEKTVVHTSQTVPDPOGSLPMGAKTPEGA 173

DB 130 DLRTGKVIQTKVDGKEVPGVIGETAKPEVGQ 162

RESULT 4

US-09-978-825-17039
; Sequence 17039, Application US/09978825
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Derrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglas, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: US/09/978,825
; CURRENT FILING DATE: 2003-01-29
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 17039
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Propionibacterium acnes
US-09-978-825-17039

Query Match 10.4%; Score 95; DB 5; Length 163;
Best Local Similarity 26.1%; Pred. No. 0.18;
Matches 40; Conservative 30; Mismatches 47; Indels 36; Gaps 9;

QY 49 PANALPARESEAAVQPVIGISQVRNMSKEKDLGTLGYVLGITMVIITIAIGAGIITL 106
DB 18 PQQARTAAARPTSLRPVLRALPQVR-----QQQWGRLGFAI--LIVVLAAGLAGLIV 69
QY 107 -----GYSYKRGKDLKE-----QHDQKVCEREMORTPL-----LSAFTNPTCEIV 147
DB 70 LNTTIOAQSMQIAQETRNINLVLQHQAVALAABVDHLRGQNLQEQAKKLGMRPNPYGSYI 129
QY 148 DEKTVVHTSQTVPDPOGSLPMGAKTPEGA 173
DB 130 DLRTGKVIQTKVDGKEVPGVIGETAKPEVGQ 162

RESULT 5

US-10-057-498-17039
; Sequence 17039, Application US/10057498
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Derrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglas, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
; FILE REFERENCE: 210121.514
; CURRENT APPLICATION NUMBER: US/10/057,498
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 29212
; SEQ ID NO 17039
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Propionibacterium acnes
US-10-057-498-17039

Query Match 10.4%; Score 95; DB 6; Length 163;
Best Local Similarity 26.1%; Pred. No. 0.18;
Matches 40; Conservative 30; Mismatches 47; Indels 36; Gaps 9;

QY 49 PANALPARESEAAVQPVIGISQVRNMSKEKDLGTLGYVLGITMVIITIAIGAGIITL 106


```

Db 18 PQQARTAAAPTSLRPVLRALPQVR-----QQMGRIGFAI---LIVVMIAAGLAGLIV 69
Qy 107 -----SYSYRGKDLKE-----QHDQVCEREMORITLPL-----LSAFTNPTCEIV 147
Db 70 LNTTIAQNSQIAQINQVRLNVLVQHQAVALAAVDHLRBPQLQRAKKGMRNPNPGSYI 129
Qy 148 DEKTVVHTISQTVDPQF-----GST--PLMQ 173
Db 130 DLRTGVIGVQTKVDKEVPGVIGETARPEVQ 162

RESULT 6
US-10-369-493-10111
; Sequence 10111, Application US/10369493
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 10111
; LENGTH: 850
; TYPE: PRT
; ORGANISM: magnetite-containing magnetic coccus
US-10-369-493-10111

Query Match 9.6%; Score 87.5; DB 6; Length 850;
Best Local Similarity 22.2%; Pred. No. 7.8;
Matches 41; Conservative 18; Mismatches 73; Indels 53; Gaps 4;

Qy 11 PCEDLRCPEFTTSQALPAFTTE-----IQEASGPGADEVQVAPANALPARSEAAAVQPV- 65
Db 72 PAVVAAPAAAPAVVAAPVVAQPAPEAPVAVEEVEKPELPAAAPAAAPAPVR 131
Qy 66 -----IGISQVRNMSKEKDLGTLGYVLGTMVITIIAGGIIIGSYR 112
Db 132 LLMQVVEEPEKRLKLSAQREEMARKRTEDVLS-----RR 165
Qy 113 GKDLKEQHDQVCEREMORITLPLSAFTNPTCEIVDEKTVVHTISQTVDPQEGSTPLMG 172
Db 166 LNLQLEFLREQKREDRAKEAVALAK-----KEKPVAAATVAAAALVAVGRTPRED 216
Qy 173 QAGTP 177
Db 217 SAGEP 221

RESULT 7
US-10-289-762-174
; Sequence 174, Application US/10289762
; GENERAL INFORMATION:
; APPLICANT: Grifais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/10/289,762
; CURRENT FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 174
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: SITE

```

[illegible]

Mon Apr 7 10:23:04 2003

us-10-057-951-2_copy_85_263.rapn

Page 4

Qy	20	TTSCALFAFTTELOE-----	-ASEGGADEVQFAPNAPARESEAA-	VQPIGIGS	69
Db	235	TOEGSLFAFTTRDITLLGVSFLVIAPEHDDDSIV-----	SEQRDEVAVAYQESLRKS		288
Qy	70	ORVAMNS-KKKKDLGTLGY----	VLGITMYV-----	IIAIGAGIILIGYGYKKRGDLKQ	119
Db	289	ERDRISSVKTKTGTFPGNVAKHPITGNLLPWIISDYVVLGGTGVVWG-----	PA		339
Qy	120	HDQVCEREMQRTLPPLSAFTNPCEIYDEKTVVHNS	157		
Db	340	HDER--DREPAEM-----	FLPIHEVAVDNGVCISHN	369	

```

RESULT 9
US-09-701-271A-4
Sequence 4, Application US/09701271A
GENERAL INFORMATION:
APPLICANT: Ruelle, Jean-Louis
APPLICANT: Tommassen, Johannes Petrus Maria
TITLE OF INVENTION: Neisseria Meningitidis Antigenic
TITLE OF INVENTION: Polypeptides, Corresponding Polynucleotides and Protectives
TITLE OF INVENTION: Antibodies
FILE REFERENCE: BM45323
CURRENT APPLICATION NUMBER: US/09/701,271A
CURRENT FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: PCT/EP99/03603
PRIOR FILING DATE: 1999-05-26
PRIOR APPLICATION NUMBER: 9811260.0
PRIOR FILING DATE: 1998-05-26
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 769
TYPE: PR1
ORGANISM: Neisseria meningitidis
US-09-701-271A-4

```

```

Query Match: 8.8%; Score 80.5; DB 5; Length 769;
Best Local Similarity 24.4%; Pzed. No. 33;
Matches 49; Conservative 26; Mismatches 69; Indels 57; Gaps 11.

Cy 3 EAGVPEKPCEDLACRPETTS--QALAFITTEIQEAS-----EGRPADENVQFAP-- 50
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 150 QAAAPSTKSAVSVSKPETPAKQAAAPFTESVSVSAPSPAKQAAASAOQAAAPAKQ 209
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cy 51 -NALPASEAAAVQVPIGVSQRVRMNSKERKDLGTLGYVLGITMMVIIAIGAGIILGVS 109
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 210 QTAAPAAQAAAPAAKQNTIDFR-----KQGNAG-----IIEIAA---LGFA 248
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cy 110 YKRGKDLKEHDDOKVCEEREMORTITLPLSAFTNPT-----CEIVDEKTVV--VHTSOTPV 162
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 249 GQ--FDISQGDH-----HIVTLKQHTLPTTLQRLSDVADPKPVQCVTLKRLNND 297
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cy 163 PQ-----EGSNPLMGAGATPG 178
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 298 TQLITTAGNWEVYNKSAAPG 318

RESULT 10
US-09-724-676-63523
; Sequence 63523, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ. ID NOS.: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 63523
;
; LENGTH: 345
; TYPE: PRT f
;

```

```

: ORGANISM Homo sapiens
US-09-724-676-63523

Query Match      8 7%; Score 60; DB 5; Length 345;
Best Local Similarity 22.4%; Pred. No. 13;
Matches 55; Conservative 20; Mismatches 77; Indels 94; Gaps 11.

QY      2  GEAGV--PERRPC-----EDLR-----PETS-----QALPFTTEIQ 33
DB      54  GECAGPEPPTPLPSGHLNNNCARLTHFRNDHVPQGTVGACISGIRSLPATRAVAR 113
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      34  E-----ASGSPADVEQV-FAPAPALPARK--EAAAVQPIPTGSCQVRRNNSREKD 81
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      114  DRLVLCDRASSGAGAVEAVAFSPARDLPDSSLIQAAHAIVAAITOR----- 163
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      82  LGTLGVLYGIVTMMVITIIAIGAGIIIG-----YSYRKGLDKE 118
DB      164  -GNSSLDLATVEKAVETVVTGSSSTGLVPLVCGAFSLMILACVLCVMMTRRRK----- 218
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      119  OHDKVCEREMQKIT-----LPISAFTNPICEIYDEKTVVHTSQTPEVDPQEGSTP 169
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      219  -----ERERSRLPRESANNQMAPLPIRNPPIERPGGHQDVLVYCKCKFTPPBRADA 271
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      170  LMGAG 175
        : : : : :
DB      272  LPPAG 277

```

```

RESULT 11
US-09-724-676A-63523
; Sequence 63523, Application US/09724676A
GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OR INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 63523
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-63523

```

```

Query Match 8.7%; Score 80; DB 5; Length 345;
Best Local Similarity 22.4%; Pred. No. 13;
Matches 55; Conservative 20; Mismatches 77; Indels 94; Gaps 11.

QY      2  GAGCV--PERRPC-----EDLRC-----PETS-----QALPFTTEIQ 33
Db      54  GCGCAEPPSTPCIPSGHLDNNCCARLTLEHNRDHPQGTVGACISGISLTPATRAVAR 113
QY      34  E-----ASGCPADAEVQ--PAPANAIPARS--EAAAQPIGISCRVRNNSREKXD 81
Db      114  DRLIYLICDRASGASAVEAVSFGARDLPDSSLICGAHAIVAAITQ-----163
QY      82  LGTLGYVLGITMVIITIAIGAGITIG-----YSYRGKDLKE 118
Db      164  GNSLLILATEYKVEFTVVTGSSSTGLLVPLYCGAFSVLMCLAVLCVMWTRRRK----218
QY      119  QHDKVCEMRQKIT-----LPLSAFTNPCEIYDEKTVVHNHSQCPVPOEGSTP 169
Db      219  -----ERRSRKLPEESANNQMAPLPINPIERPGCHDVLVYQCKNTPTPPPRADA 271
QY      170  LMGQAG 175
Db      272  LPGPAG 277

RESULT 12
US-09-724-676-63525
; Sequence 63525, Application US/09724676

```

```

; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 63525
; LENGTH: 1226
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (561)..(561)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (565)..(565)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-09-724-676-63525
```

```

Query Match
Best Local Similarity 22.4%; Score 80; DB 5; Length 1226;
Matches 55; Conservative 20; Mismatches 77; Indels 94; Gaps 11;
```

```

QY 2 GEGAV--PEKPPC-----EDLRC-----PETTS-----QALPATTETIQ 33
DB 935 GEGAEPPSTPCIPRSGHLDNNCARLTHFNRDHVPQGTIVGALCSGIRSLPATRAVAR 994
QY 34 E-----ASEGPADEVQV--FAPANALPARS--EAAAVQPIGIGSQVRNMSKEKDD 81
DB 995 DRLVLVLCRRASSGASAVEVAVSFSPARDLPDSSLIGAAHAIVAALTQR----- 1044
QY 82 LGTLGVVIGITMWTIIAIGAGIILG-----YSYRGKDLKE 118
DB 1045 -GNSSLLAVTEVAVETVVTGSSSTGLVPLVLCGAFSVLMACVYLCAVWTRKRRK--- 1099
QY 119 QHDQVCEREMQRT-----LPLSAFTNPCEIVDEKTVVHTSQTVPDPOEGSTP 169
DB 1100 -----ERERSRLPREBSANNQMAPLNPINPIERPGHKDVLVYOCKNFTPPRADA 1152
QY 170 LMGQAG 175
DB 1153 LPPGAG 1158
```

```

RESULT 13
US-09-724-676A-63525
; Sequence 63525, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 63525
; LENGTH: 1226
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (561)..(561)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (565)..(565)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-09-724-676A-63525
```

Query Match 8.7%; Score 80; DB 5; Length 1226;

```

Best Local Similarity 22.4%; Pred. No. 66;
Matches 55; Conservative 20; Mismatches 77; Indels 94; Gaps 11;

QY 2 GEGAV--PEKPPC-----EDLRC-----PETTS-----QALPATTETIQ 33
DB 935 GEGAEPPSTPCIPRSGHLDNNCARLTHFNRDHVPQGTIVGALCSGIRSLPATRAVAR 994
QY 34 E-----ASEGPADEVQV--FAPANALPARS--EAAAVQPIGIGSQVRNMSKEKDD 81
DB 995 DRLVLVLCRRASSGASAVEVAVSFSPARDLPDSSLIGAAHAIVAALTQR----- 1044
QY 82 LGTLGVVIGITMWTIIAIGAGIILG-----YSYRGKDLKE 118
DB 1045 -GNSSLLAVTEVAVETVVTGSSSTGLVPLVLCGAFSVLMACVYLCAVWTRKRRK--- 1099
QY 119 QHDQVCEREMQRT-----LPLSAFTNPCEIVDEKTVVHTSQTVPDPOEGSTP 169
DB 1100 -----ERERSRLPREBSANNQMAPLNPINPIERPGHKDVLVYOCKNFTPPRADA 1152
QY 170 LMGQAG 175
DB 1153 LPPGAG 1158
```

```

RESULT 14
US-09-724-676-63524
; Sequence 63524, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 63524
; LENGTH: 1264
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (599)..(599)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (603)..(603)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-09-724-676-63524
```

```

Query Match
Best Local Similarity 22.4%; Score 80; DB 5; Length 1264;
Matches 55; Conservative 20; Mismatches 77; Indels 94; Gaps 11;
```

```

QY 2 GEGAV--PEKPPC-----EDLRC-----PETTS-----QALPATTETIQ 33
DB 973 GEGAEPPSTPCIPRSGHLDNNCARLTHFNRDHVPQGTIVGALCSGIRSLPATRAVAR 1032
QY 34 E-----ASEGPADEVQV--FAPANALPARS--EAAAVQPIGIGSQVRNMSKEKDD 81
DB 1033 DRLVLVLCRRASSGASAVEVAVSFSPARDLPDSSLIGAAHAIVAALTQR----- 1082
QY 82 LGTLGVVIGITMWTIIAIGAGIILG-----YSYRGKDLKE 118
DB 1083 -GNSSLLAVTEVAVETVVTGSSSTGLVPLVLCGAFSVLMACVYLCAVWTRKRRK--- 1137
QY 119 QHDQVCEREMQRT-----LPLSAFTNPCEIVDEKTVVHTSQTVPDPOEGSTP 169
DB 1138 -----ERERSRLPREBSANNQMAPLNPINPIERPGHKDVLVYOCKNFTPPRADA 1190
QY 170 LMGQAG 175
DB 1191 LPPGAG 1196
```

```

RESULT 15
US-09-724-676A-63524
; Sequence 63524, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724, 676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 63524
; LENGTH: 1264
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (599)..(599)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (603)..(603)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-09-724-676A-63524

```

Query Match 8.7%; Score 80; DB 5; Length 1264;

Best Local Similarity 22.4%; Pred. No. 69; Mismatches 77; Indels 94; Gaps 11;

Matches 55; Conservative 20; Mismatches 77; Indels 94; Gaps 11;

```

QY 2 GEAGV--PEKRP-----EDLRC-----PETTS-----QALPATTETIQ 33
Db 973 GEGCAEPSPSTCLPRSGHLNNCARLTLEHNDHVPQGTWGAICSGIRSLPATRAVAR 1032
QY 34 E-----ASEGQADEVOV--FAPANALPARS--EAAAVQVYIGISQEVNMSKXKD 81
Db 1033 DRLVLICDRASSGASAVEVAVSFPARDLPDSLIQGAHAIVAAITQR----- 1082
QY 82 LGTGVVLGITMMVITIIAIGAILG-----YSYRKGDLKE 118
Db 1083 -GNSSLIAVTEVKVETVTGSGTGLVPLCGAPSVMLACVLCVWMTTRRRK---- 1137
QY 119 QHOKVCEBEMQRI-----LPSATNPTCEIVDEKTVVYHTSQTPVDQEGSTP 169
Db 1138 -----ERRSRLPRESANNQWAPLPIRNPIERPGGHKDVLYQCKNFTPPPRADDA 1190
QY 170 LMGQAG 175
Db 1191 LPGPAG 1196

```

Search completed: April 7, 2003, 09:30:51
Job time : 29.7942 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2003, 09:16:28 ; Search time 17.8628 Seconds
(without alignments)
294.842 Million cell updates/sec

Title: US-10-057-951-2_COPY_85_263

Perfect score: 916
Sequence: 1 SGEAGVPEKRPEDLRCPET.....PVDQEGSTPLMGQAGTPGA 179

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTCUTS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	916	100.0	263	4	US-09-411-977-2
2	80	8.7	1212	4	US-09-214-278-3
3	80	8.7	1238	4	US-09-214-278-5
4	77.5	8.5	463	1	US-09-134-001C-3973
5	77.5	8.5	595	1	US-08-225-989-2
6	77.5	8.5	595	1	US-08-570-923-2
7	77.5	8.5	595	1	US-08-580-014-2
8	77.5	8.5	595	4	US-09-079-785-2
9	75.5	8.2	219	4	US-09-134-001C-5651
10	75	8.2	595	2	US-08-232-087A-2
11	75	8.2	595	4	US-09-006-353A-9
12	75	8.2	595	4	US-09-573-986-9
13	74.5	8.1	587	1	US-08-398-008A-2
14	74.5	8.1	587	1	US-08-893-333-2
15	74	8.1	1148	4	US-08-882-046-4
16	72.5	7.9	225	4	US-09-134-001C-3215
17	72.5	7.9	400	1	US-08-351-473B-5
18	72.5	7.9	400	4	US-08-450-962-4
19	72.5	7.9	400	4	US-08-450-962-6
20	72	7.9	1780	1	US-08-769-309A-5
21	72	7.9	1780	1	US-08-769-309A-5
22	71.5	7.8	707	4	US-08-994-570-5
23	71	7.8	652	4	US-09-228-986-80
24	70.5	7.7	1182	4	US-09-287-354-6
25	69.5	7.6	1025	3	US-08-304-309-4
26	69.5	7.6	1025	3	US-08-991-942-4
27	69	7.5	1025	2	US-08-304-309-2

28	69	7.5	1025	3	US-08-991-942-2	Sequence 2, Appl1
29	69	7.5	1025	4	US-09-138-103-2	Sequence 2, Appl1
30	69	7.5	1025	5	PCT-US95-04567-4	Sequence 4, Appl1
31	68.5	7.5	673	2	US-08-455-073A-6	Sequence 6, Appl1
32	68.5	7.5	683	2	US-08-477-396A-17	Sequence 17, Appl1
33	68	7.4	1257	3	US-08-611-729A-8	Sequence 8, Appl1
34	67.5	7.4	683	1	US-07-878-960-2	Sequence 2, Appl1
35	67	7.3	271	4	US-09-414-436-1	Sequence 1, Appl1
36	67	7.3	388	1	US-08-087-772A-2	Sequence 2, Appl1
37	67	7.3	458	4	US-09-134-001C-4663	Sequence 4, Appl1
38	67	7.3	531	4	US-09-134-001C-4920	Sequence 4663, Ap
39	66.5	7.3	400	1	US-08-351-473B-4	Sequence 4920, Ap
40	66.5	7.3	1248	4	US-08-882-046-6	Sequence 4, Appl1
41	66.5	7.3	10182	4	US-09-134-001C-3159	Sequence 6, Appl1
42	66	7.2	279	4	US-08-339-214-24	Sequence 3159, Ap
43	66	7.2	279	4	US-08-339-214-32	Sequence 24, Appl1
44	66	7.2	427	4	US-09-086-483A-4	Sequence 32, Appl1
45	66	7.2	427	4	US-09-041-886-2	Sequence 4, Appl1

ALIGNMENTS

```
RESULT 1
US-09-411-977-2
; Sequence 2, Application US/09411977
; Patent No. 6372473
; GENERAL INFORMATION:
; APPLICANT: Moore, Paul A.
; APPLICANT: Ruben, Steven M.
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
; FILE REFERENCE: Paj78P1
; CURRENT APPLICATION NUMBER: US/09/411,977
; CURRENT FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: 09/084,491
; EARLIER FILING DATE: 1998-05-27
; EARLIER APPLICATION NUMBER: 60/048,000
; EARLIER FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-411-977-2

Query Match      100.0%; Score 916; DB 4; Length 263;
Best Local Similarity 100.0%; Pred. No. 8.1e-98;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGEAGVPEKRPEDLRCPETTSOALPAFTTBIOBASGPGADEQVAPANALPARSEAA 60
DB 85 SGEAGVPEKRPEDLRCPETTSOALPAFTTBIOBASGPGADEQVAPANALPARSEAA 144

QY 61 AVQPIGISQVRNMSKKKDLGTLGYVLGITMVVITIAIGAGIILGYSYKRGDLKEQH 120
DB 145 AVQPIGISQVRNMSKKKDLGTLGYVLGITMVVITIAIGAGIILGYSYKRGDLKEQH 204

QY 121 DQKVEREMORTLPLSFTNPTCEIYDEKVVTHTSOTVPDPEGSTPLMGAGTPGA 179
DB 205 DQKVEREMORTLPLSFTNPTCEIYDEKVVTHTSOTVPDPEGSTPLMGAGTPGA 263

RESULT 2
US-09-214-278-3
; Sequence 3, Application US/09214278
; Patent No. 6291210
; GENERAL INFORMATION:
; APPLICANT: Sakano, Seiji
; APPLICANT: Itoh, Akira
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-8576
```

;; CURRENT APPLICATION NUMBER: US/09/214,278
;; CURRENT FILING DATE: 1999-01-26
;; NUMBER OF SEQ ID NOS: 32
;; SOFTWARE: Patentin Ver. 2.1
;; SEQ ID NO 3
;; LENGTH: 1212
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-214-278-3

Query Match 8.7%; Score 80; DB 4; Length 1212;
Best Local Similarity 22.4%; Pred. No. 3.4; Indels 94; Gaps 11;
Matches 55; Conservative 20; Mismatches 77;

QY 2 GEAGV--PEKRPC-----EDLRC-----PETTS-----QALPFTTEIQ 33
DB 921 GEGCAEPPSTPCLPBSGHLNNCARLTHFNNDHVPQGTGALCSGIRSLPATRAVAR 980
QY 34 E-----ASBGPADDEVQV--FAPANMLPARS--EAAAVPVGISQRYRMSKEKDD 81
DB 981 DRLIVLCDBRASSGASAVEVAVSFSPARDLPDSSLIOGAHAIVAAITQR----- 1030
QY 82 LGTLGVVLGITMMVITIAIGAGIILG-----YSYKRGKDKE 118
DB 1031 -GNSSLILAVTEVKVETVVGSSSTGLVPLVCGAFSVMLACVLCVWMTKRK----- 1085
QY 119 QHDKVCVEREMQRT-----LPLSAFTNPTCEIVDEKTVVHTSOTPVDPQEGSTP 169
DB 1086 -----ERRRSRLPREESANNOMAPLNPRIERPGRGHKQVLYOCKNFPPPRRADEA 1138
QY 170 IMGQAG 175
DB 1139 LPEPAG 1144

RESULT 3
US-09-214-278-5
;; Sequence 5, Application US/09214278
;; Patent No. 6291210
;; GENERAL INFORMATION:
;; APPLICANT: Itoh, Akira
;; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
;; FILE REFERENCE: KP-8576
;; CURRENT APPLICATION NUMBER: US/09/214,278
;; CURRENT FILING DATE: 1999-01-26
;; NUMBER OF SEQ ID NOS: 32
;; SOFTWARE: Patentin Ver. 2.1
;; SEQ ID NO 5
;; LENGTH: 1238
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-214-278-5

Query Match 8.7%; Score 80; DB 4; Length 1238;
Best Local Similarity 22.4%; Pred. No. 3.5; Indels 94; Gaps 11;
Matches 55; Conservative 20; Mismatches 77;

QY 2 GEAGV--PEKRPC-----EDLRC-----PETTS-----QALPFTTEIQ 33
DB 947 GEGCAEPPSTPCLPBSGHLNNCARLTHFNNDHVPQGTGALCSGIRSLPATRAVAR 1006
QY 34 E-----ASBGPADDEVQV--FAPANMLPARS--EAAAVPVGISQRYRMSKEKDD 81
DB 1007 DRLIVLCDBRASSGASAVEVAVSFSPARDLPDSSLIOGAHAIVAAITQR----- 1056
QY 82 LGTLGVVLGITMMVITIAIGAGIILG-----YSYKRGKDKE 118
DB 1057 -GNSSLILAVTEVKVETVVGSSSTGLVPLVCGAFSVMLACVLCVWMTKRK----- 1111
QY 119 QHDKVCVEREMQRT-----LPLSAFTNPTCEIVDEKTVVHTSOTPVDPQEGSTP 169
DB 1112 -----ERRRSRLPREESANNOMAPLNPRIERPGRGHKQVLYOCKNFPPPRRADEA 1164

QY 170 IMGQAG 175
DB 1165 LPEPAG 1170

RESULT 4
US-09-134-001C-3973
;; Sequence 3973, Application US/09134001C
;; Patent No. 6180370
;; GENERAL INFORMATION:
;; APPLICANT: Lynn Doucette-Stamm et al
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;; FILE REFERENCE: GTC-007
;; CURRENT APPLICATION NUMBER: US/09/134,001C
;; CURRENT FILING DATE: 1998-08-13
;; PRIOR APPLICATION NUMBER: US 60/064,964
;; PRIOR FILING DATE: 1997-11-08
;; PRIOR FILING DATE: 1997-08-14
;; NUMBER OF SEQ ID NOS: 5674
;; SEQ ID NO 3973
;; LENGTH: 463
;; TYPE: PRT
;; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3973

Query Match 8.5%; Score 77.5; DB 4; Length 463;
Best Local Similarity 31.1%; Pred. No. 1.6;
Matches 28; Conservative 14; Mismatches 37; Indels 11; Gaps 3;

QY 66 IGISQR--VRMSKEKDLGTLGVVLGITMMVITIA--IGAGIILGYGKRGKDKEHQD 122
DB 116 IGIRERQILMLNDRDNTSGIVKLIETVITIFIEIGALLAFYRDNPDLE----- 170
QY 123 KVCEREMQRTPLPLSAFTNPTCEIVDEKTV 152
DB 171 ---NALMGIFVSVSATNNGIDITGESLV 197

RESULT 5
US-08-225-989-2
;; Sequence 2, Application US/08225989
;; Patent No. 5480981
;; GENERAL INFORMATION:
;; APPLICANT: Goodwin, Raymond G.
;; APPLICANT: Smith, Craig A.
;; APPLICANT: Armitage, Richard J.
;; TITLE OF INVENTION: Gruse, Hans-Jurgen
;; NUMBER OF SEQUENCES: 23
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Kathryn A. Seese, Immunex Corporation
;; STREET: 51 University Street
;; CITY: Seattle
;; STATE: Washington
;; COUNTRY: USA
;; ZIP: 98101
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: Apple Macintosh
;; OPERATING SYSTEM: Apple 7.1
;; SOFTWARE: Microsoft Word, Version 5.1a
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/225,989
;; FILING DATE: 12 April 1994
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/966,775
;; FILING DATE: 27-OCT-1992
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 907,224
FILING DATE: 01-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 899,660
FILING DATE: 15-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 892,459
FILING DATE: 02-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 889,717
FILING DATE: 26-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Seese, Kathryn A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2804-E
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 595 amino acids
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-225-989-2

Query Match 8.5%; Score 77.5; DB 1; Length 595;
Best Local Similarity 23.6%; Pred. No. 2.3;
Matches 46; Conservative 32; Mismatches 74; Indels 43; Gaps 11;

QY 4 AGVPEKPCEDLRCEPTTSQALPAFT---TEIQASGPGADVQVFAPAMLPARSEA 59
DB 307 AGETVTKP-QDMAKQDTTFEAPLGTQPCNPFPNGEAPAST-----SPTQSLVDSQA 360
QY 60 AAVQPVIGISQVRNMSKEDLGTGLGVLTMMVITIIAGIILIGYSYKRGKDLKEQ 119
DB 361 SKTLP-IPTSAPVALSSTGKPEVLA-GPVLFWVILVIVVSSAFL-----L 406
QY 120 HDQKVCEREMQ---ITPLSAFTNPCEIVDEKTVVATSTQ-----PYDPQEG 166
DB 407 CHRRACKRIRQKLHCYPVQT-SQPLLELVDSRP---RRSSTQLRSGASVTEPVAEREG 462
QY 167 --STPLMGQAGTPGA 179
DB 463 LMSQPLMETCHSVGA 477

RESULT 6
US-08-570-923-2
Sequence 2, Application US/08570923
Patent No. 5677430
GENERAL INFORMATION:
APPLICANT: Goodwin, Raymond G.
APPLICANT: Smith, Craig A.
APPLICANT: Armitage, Richard J.
APPLICANT: Gruss, Hans-Jurgen
TITLE OF INVENTION: No. 5677430el Cytokine That Binds CD30
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kathryn A. Seese, Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple 7.1
SOFTWARE: Microsoft Word, Version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/570,923

FILING DATE: 12-DEC-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/225,989
FILING DATE: 12 APRIL 1994
APPLICATION NUMBER: US 07/966,775
FILING DATE: 27-OCT-1992
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 907,224
FILING DATE: 01-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 899,660
FILING DATE: 15-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 892,459
FILING DATE: 02-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 889,717
FILING DATE: 26-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Seese, Kathryn A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2804-E
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 595 amino acids
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-570-923-2

Query Match 8.5%; Score 77.5; DB 1; Length 595;
Best Local Similarity 23.6%; Pred. No. 2.3;
Matches 46; Conservative 32; Mismatches 74; Indels 43; Gaps 11;

QY 4 AGVPEKPCEDLRCEPTTSQALPAFT---TEIQASGPGADVQVFAPAMLPARSEA 59
DB 307 AGETVTKP-QDMAKQDTTFEAPLGTQPCNPFPNGEAPAST-----SPTQSLVDSQA 360
QY 60 AAVQPVIGISQVRNMSKEDLGTGLGVLTMMVITIIAGIILIGYSYKRGKDLKEQ 119
DB 361 SKTLP-IPTSAPVALSSTGKPEVLA-GPVLFWVILVIVVSSAFL-----L 406
QY 120 HDQKVCEREMQ---ITPLSAFTNPCEIVDEKTVVATSTQ-----PYDPQEG 166
DB 407 CHRRACKRIRQKLHCYPVQT-SQPLLELVDSRP---RRSSTQLRSGASVTEPVAEREG 462
QY 167 --STPLMGQAGTPGA 179
DB 463 LMSQPLMETCHSVGA 477

RESULT 7
US-08-580-014-2
Sequence 2, Application US/08580014
Patent No. 5753203
GENERAL INFORMATION:
APPLICANT: Goodwin, Raymond G.
APPLICANT: Smith, Craig A.
APPLICANT: Armitage, Richard J.
APPLICANT: Gruss, Hans-Jurgen
TITLE OF INVENTION: No. 5753203el Cytokine That Binds CD30
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kathryn A. Seese, Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington

```

; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple 7.1
; SOFTWARE: Microsoft Word, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/580,014
; FILING DATE: 20-DEC-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/225,989
; FILING DATE: 12 APRIL 1994
; APPLICATION NUMBER: US 07/966,775
; FILING DATE: 27-OCT-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 907,224
; FILING DATE: 01-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 899,660
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 892,459
; FILING DATE: 02-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 889,717
; FILING DATE: 26-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Seese, Kathryn A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2804-E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)233-0644
; TELEX: 756822
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 595 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-580-014-2

Query Match      8.5%; Score 77.5; DB 1; Length 595;
Best Local Similarity 23.6%; Pred. No. 2.3;
Matches 46; Conservative 32; Mismatches 74; Indels 43; Gaps 11;

QY 4 AGVEKPCEDLCPEETSQALPAFT---TEIQASGPGADDEVQVAPANALPARSEA 59
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 307 AGETVTKP-QDMAEKDTFEAPPLGTQPCNPTPENGEAPAST-----SPTQSLVDSQA 360
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

QY 60 AAVQPVIGISQVRNNSKKKDLGTLGYLITMVTIIAIGAGIILGYSYKRGDKLEQ 119
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 361 SKTLP-IPTSAFVALSTGKPVLDAGPVLFWIIVLVVVGSSAFL-----L 406
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

QY 120 HDQKVCEREMOR---ITPLSAFTNPTCEIYDEKTVVHTSQT-----PVDPQEG 166
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 407 CHRACRKRIRQKHLCTYPVQT-SQPKELVDSRP---RRSSTQLRSGASVTEPVAEBRG 462
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

QY 167 --STPLMGOAGTPGA 179
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 463 LMSQPLMETCHSVGA 477
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

RESULT 8
US-09-079-785-2
; Sequence 2, Application US/09079785
; Patent No. 6143869
; GENERAL INFORMATION:
; APPLICANT: Goodwin, Raymond G.
; APPLICANT: Smith, Craig A.
```

```

; APPLICANT: Armitage, Richard J.
; APPLICANT: Gruss, Hans-Jurgen
; TITLE OF INVENTION: No. 6143869el Cytokine That Binds CD30
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kathryn A. Seese, Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple 7.1
; SOFTWARE: Microsoft Word, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/079,785
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/225,989
; FILING DATE: 12 APRIL 1994
; APPLICATION NUMBER: US 07/966,775
; FILING DATE: 27-OCT-1992
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 907,224
; FILING DATE: 01-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 899,660
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 892,459
; FILING DATE: 02-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 889,717
; FILING DATE: 26-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Seese, Kathryn A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2804-E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)233-0644
; TELEX: 756822
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 595 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-079-785-2

Query Match      8.5%; Score 77.5; DB 4; Length 595;
Best Local Similarity 23.6%; Pred. No. 2.3;
Matches 46; Conservative 32; Mismatches 74; Indels 43; Gaps 11;

QY 4 AGVEKPCEDLCPEETSQALPAFT---TEIQASGPGADDEVQVAPANALPARSEA 59
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 307 AGETVTKP-QDMAEKDTFEAPPLGTQPCNPTPENGEAPAST-----SPTQSLVDSQA 360
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

QY 60 AAVQPVIGISQVRNNSKKKDLGTLGYLITMVTIIAIGAGIILGYSYKRGDKLEQ 119
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 361 SKTLP-IPTSAFVALSTGKPVLDAGPVLFWIIVLVVVGSSAFL-----L 406
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

QY 120 HDQKVCEREMOR---ITPLSAFTNPTCEIYDEKTVVHTSQT-----PVDPQEG 166
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 407 CHRACRKRIRQKHLCTYPVQT-SQPKELVDSRP---RRSSTQLRSGASVTEPVAEBRG 462
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

QY 167 --STPLMGOAGTPGA 179
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB 463 LMSQPLMETCHSVGA 477
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
```



```

RESULT 9
US-09-134-001C-5651
; Sequence 5651, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5651
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5651

Query Match      8.2%; Score 75.5; DB 4; Length 219;
Best Local Similarity 26.7%; Pred. No. 0.88;
Matches 32; Conservative 22; Mismatches 33; Indels 33; Gaps 7;

QY 75 NSKEKDL-GTLGVLTGTMVILIAAGILIGSY-----KRG-----113
DB 7 NREVNNSGIGSKFISAIIVLLILIGLAFGI--YGFVDSKGNERLSDKTTQOEKKD 63

QY 114 -KDLKEQHDQVRCRE---MORTLPLSAFTNPTCEIYDEKTVVHTSQRPV--DPQEG 166
DB 64 DKDKKKDKXSVBEKKNNTQQTQGVPTQTQQTQQTQV--OTPRPPTTQTIPVKRNPQTG 121

RESULT 10
US-08-232-087A-2
; Sequence 2, Application US/08232087A
; Patent No. 5866372
; GENERAL INFORMATION:
; APPLICANT: Stein, Harald
; APPLICANT: D Rkop, Horst
; APPLICANT: Latza, Ute
; TITLE OF INVENTION: Lymphoid CD30-Antigen
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,087A
; FILING DATE: 08-SEP-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 756-103P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 203-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 595 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-232-087A-2

Query Match      8.2%; Score 75; DB 2; Length 595;
Best Local Similarity 23.1%; Pred. No. 4.4;
Matches 43; Conservative 31; Mismatches 70; Indels 42; Gaps 10;

QY 13 EDLRCPPTSQALPAFT---TEIQASGPGADVQVFAFANMLPARSAAVQPIGTI 68
DB 315 QDMVKQTTTEAPLGTQDPNPTENGEPAST-----SPQSILVDSQASKTLP-IPR 368

QY 69 SQRYMNSKEKDLGTLGVLTGTMVILIAAGILIGSYKKGKDLKEQHDQVRCRE 128
DB 369 SAPVALSTGKPLVLA-GVLFVWILVIVVVGSSAFL-----LCHRRACRRK 415

QY 129 MQR---ITLPLSAFTNPTCEIYDEKTVVHTSQT-----PVPQEG--STPLMGQ 173
DB 416 IROKLHLCEYVQT-SQPKLELVDSRP---RRSTQLRGSASVTEPVAEERGLMSQPLMET 471

QY 174 AGTPGA 179
DB 472 CHSVGA 477

RESULT 11
US-09-006-353A-9
; Sequence 9, Application US/09006353A
; Patent No. 6261801
; GENERAL INFORMATION:
; APPLICANT: WEI, YING-PEI
; APPLICANT: YU, GHO-LIANG
; APPLICANT: GENTZ, REINER
; APPLICANT: RUBEN, STEVEN
; TITLE OF INVENTION: TUMOR NECROSIS FACTOR RECEPTOR 5
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/006,353A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF341
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 595 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-006-353A-9

Query Match      8.2%; Score 75; DB 4; Length 595;
Best Local Similarity 23.1%; Pred. No. 4.4;
Matches 43; Conservative 31; Mismatches 70; Indels 42; Gaps 10;

```

```

RESULT 12
US-09-573-986-9
: Sequence 9, Application US/09573986
: Patent No. 6455040
:
GENERAL INFORMATION:
:
APPLICANT: Wei, Ying-Fei
APPLICANT: Ni, Jian
APPLICANT: Gentz, Reiner
APPLICANT: Ruben, Steven
:
TITLE OF INVENTION: Tumor Necrosis Factor Receptor 5
:
FILE REFERENCE: 1488.1280004
:
CURRENT APPLICATION NUMBER: US/09/573,986
:
CURRENT FILING DATE: 2000-05-18
:
NUMBER OF SEQ ID NOS: 27
:
SOFTWARE: PatentIn Ver. 2.1
:
SEQ ID NO 9
:
LENGTH: 595
:
TYPE: PRT
:
ORGANISM: Homo sapiens
:
US-09-573-986-9

```

```

OY 13 EDLRCPTTSQALPAT-----TEIOABEGGCBDEVFAPNALPABEALAAQPVIGI 68
Db 315 QDMAEKOTTEAPDICTQDPCNPTEPENGAPAST-----SPYOSLIVDSQAKETLP-1PT 366
OY 69 SQRVBNRSKEKDLGTGLGYVLGITMMVITIIAIGAILGSYVRKDLKEQDOKYCERE 128
Db 369 SAPALASSTKCPULDA-GPYVLFWYLLVLVVYSSAFL-----LCHRACKKR 416
OY 129 MOR---ITLELSAFTNPCEIVDEKTVVHTSQT-----PVDPOEG--STPLMGQ 175
Db 416 IROKLHLICYPVQ--SQPKLELVDSBP-----RRSSTQOLRSAGSVTEPVAEBEGLMSQPLMET 471
OY 174 AGTRGA 179
Db 472 CHSVGA 477

```

```

1 STATE: Texas
2 ZIP: 77071
3 COUNTRY: USA
4
5 COMPUTER READABLE FORM:
6 MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
7
8 OPERATING SYSTEM: Macintosh
9 SOFTWARE: Microsoft Word 5.1
10 CURRENT APPLICATION DATA:
11 APPLICATION NUMBER: US/08/398,008A
12 FILING DATE: March 2, 1995
13 CLASSIFICATION: 435
14 PRIOR APPLICATION DATA:
15 APPLICATION NUMBER: 08/126,501
16 FILING DATE: 24-SEP-1993
17 ATTORNEY/AGENT INFORMATION:
18 NAME: Adler, Dr. Benjamin Aaron
19 REGISTRATION NUMBER: 35,423
20 REFERENCE/DOCKET NUMBER: D5705CIP
21 TELECOMMUNICATION INFORMATION:
22 TELEPHONE: (713) 777-2321
23 TELEFAX: (713) 777-6908
24
25 TELEX:
26 INFORMATION FOR SEQ ID NO: 2:
27 SEQUENCE CHARACTERISTICS:
28 LENGTH: 587
29 TYPE: amino acid
30 STRANDEDNESS:
31 TOPOLOGY: linear
32 MOLECULE TYPE: Protein
33 HYPOTHEICAL: NO
34 ANTI-SENSE: NO
35
36 US-08-398-008A-2
37
38 Query Match 8.1%; Score 74.5; DB 1; Length 587;
39 Best Local Similarity 25.2%; Pred. No. 5;
40 Matches 38; Conservative 17; Mismatches 39; Indels 57; Gaps
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
9
```

Query Match 8.1%; Score 74.5; DB 2; Length 587;
Best Local Similarity 25.2%; Pred. No. 5;
Matches 38; Conservative 17; Mismatches 39; Indels 57; Gaps 8;

QY 52 ALPARSAAVQPIVIGISORV-----RMSKEKDLGTIGVYLGITMVYII----- 97
DB 367 ALPAM-----VIGIPVYGRKHSRYEGRSKSKHKLALITG---GTTLSVIASPV 415

QY 98 ----IAIGAGIILGYST-----KRGKDLK---EHDQKVCEREMORIT 133
DB 416 AAVSVGIGVPMILAYVGVVPISLCRGGYGVSTANGKGVKIEPDEDGPIVADAMR-- 473

QY 134 LPLSAFTNPTC--EIVDEKTVVHTSQTPTVD 162
DB 474 ----ALKNPSTIGESSIGLTSVLTSGSPTD 500

RESULT 15
US-08-882-046-4
; Sequence 4, Application US/08882046
; Patent No. 6136952
; GENERAL INFORMATION:
; APPLICANT: Li, Linheng
; APPLICANT: Hood, Leroy
; APPLICANT: Krantz, Ian D.
; APPLICANT: Spinner, Nancy B.
; TITLE OF INVENTION: Human Jagged Polypeptide, Encoding
; TITLE OF INVENTION: Nucleic Acids and Methods of Use
; NUMBER OF SEQUENCES: 110
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/882,046
; FILING DATE: 25-JUN-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UW 2637
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEO ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1148 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-882-046-4

Query Match 8.1%; Score 74; DB 4; Length 1148;
Best Local Similarity 22.0%; Pred. No. 15;
Matches 54; Conservative 20; Mismatches 77; Indels 94; Gaps 11;

QY 2 GEAGV--PEKRPC-----EDLRC-----PETTS-----QALPAFTTBQ 33
DB 858 GEGCAEBPSTPCIPRSGHLDNNCARLTHFNDRDHPQGTIVAGISGIRSLPATRAVAR 917

QY 34 E-----ASEGPGADEVQV--FAPANALPARS--EAAAVQPIVIGISORVRMSKEK 81
DB 918 DRLLVLCDRASSGASAVEVAFVSFPAKDLPSLSLIGAAHAIVAAITDR----- 967

QY 82 LGTLGVYLGITMVYIIAIGAGIILG-----YSYKRGDLKE 118

DB 968 -GNSSLIAVIEVKETVVTGSSSTGLLVPLCGAFSVLMACVLCVMTTRKRRK----- 1022
QY 119 QHDQKVCEREMORIT-----LPLSAFTNPTCEIYDEKTVVHTSQTPTVDPEQSGSTP 169
DB 1023 -----ERERERLPRRESANNQMAPLNPINPIERPGHGDVLYQCKNFTPPRRRADA 1075

QY 170 IMGQA 174
DB 1076 LPGA 1080

Search completed: April 7, 2003, 09:22:33
Job time : 20.8628 secs

